

SEARCH REQUEST FORM

Requestor's

Name: _____

Serial

Number: _____

Date: _____

Phone: _____

Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 5/15/02Searcher: P. Schreiber 308-4292Terminal time: ~~308-4292~~ 8Elapsed time: 9

CPU time: _____

Total time: _____

Number of Searches: _____

Number of Databases: 13

Search Site

____ STIC

☒ CM-1 46A03

____ Pre-S

Type of Search

☒ N.A. Sequence☒ A.A. Sequence

____ Structure

____ Bibliographic

Vendors

____ IG

____ STN

____ Dialog

____ APS

____ Geninfo

____ SDC

____ DARC/Questel

☒ Other CompuLink

THIS PAGE BLANK (USPTO)

WEST

[Help](#)
[Logout](#)
[Interrupt](#)
[Main Menu](#)
[Search Form](#)
[Posting Counts](#)
[Show 8 Numbers](#)
[Edit 8 Numbers](#)
[Preferences](#)
[Cases](#)

Search Results -

Terms	Documents
l5 and l3	0

Database: US Patents Full-Text Database
US Pre-Grant Publication Full-Text Database
JPO Abstracts Database
EPO Abstracts Database
Derwent World Patents Index
IBM Technical Disclosure Bulletins

Search:

L9

[Refine Search](#)
[Recall Text](#)
[Clear](#)

Search History

DATE: Wednesday, May 15, 2002 [Printable Copy](#) [Create Case](#)

Set Name Query
side by side

Hit Count Set Name
result set

DB=USPT,PGPB,JPAB,EPAB,DWPI; PLUR=YES; OP=OR

<u>L9</u>	l5 and l3	0	<u>L9</u>
<u>L8</u>	l4 and l3	0	<u>L8</u>
<u>L7</u>	L2 and l3	1	<u>L7</u>
<u>L6</u>	l1 and l3	14	<u>L6</u>
<u>L5</u>	l1 same ny\$	60	<u>L5</u>
<u>L4</u>	l1 same nek\$	5	<u>L4</u>
<u>L3</u>	walke.in. or turner.in. or zambrowicz.in.	9697	<u>L3</u>
<u>L2</u>	L1 same (serin\$ or threoni\$)	262	<u>L2</u>
<u>L1</u>	huma\$ same kinas\$ same (recombin\$ or isolat\$ or clon\$)	2837	<u>L1</u>

THIS PAGE BLANK (USPTO)

WEST[Generate Collection](#)[Print](#)

Search Results - Record(s) 1 through 1 of 1 returned.

☐ 1. Document ID: WO 200142435 A2, AU 200125767 A

L7: Entry 1 of 1

File: DWPI

Jun 14, 2001

DERWENT-ACC-NO: 2001-381667

DERWENT-WEEK: 200140

COPYRIGHT 2002 DERWENT INFORMATION LTD

TITLE: Novel isolated human kinase polynucleotide that shares structural similarity with animal kinases including calcium/calmodulin-dependent protein kinases and serine/threonine protein kinases, useful in therapeutics

INVENTOR: ABUIN, A; DONOHO, G ; FRIEDRICH, G ; SANDS, A T ; SCOVILLE, J ; TURNER, C A ; ZAMBROWICZ, B

PRIORITY-DATA: 1999US-169428P (December 7, 1999)

PATENT-FAMILY:

PUB-NO	PUB-DATE	LANGUAGE	PAGES	MAIN-IPC
WO 200142435 A2	June 14, 2001	E	031	C12N009/00
AU 200125767 A	June 18, 2001		000	C12N009/00

INT-CL (IPC): C12 N 9/00

Full	Title	CIT.1	REV.1	CLS.1	REF.1	SEQ.1	ATT.1
CAW.1							

[Generate Collection](#)[Print](#)

Terms	Documents
L2 and I3	1

Display Format: -

[Change Format](#)[Previous Page](#)[Next Page](#)

THIS PAGE BLANK (USPTO)

66764

Schreiber, David

From: Ramirez, Delia
Sent: Tuesday, May 14, 2002 10:19 AM
To: Schreiber, David
Subject: case 09/783,320

Hi David,

I was wondering if you could do a search for me. I would like to get a standard search of seq id 4 in the nucleic and protein databases (commercial and interference).

Thank you,

Delia

Delia M. Ramirez, Ph.D.
Patent Examiner - Art Unit 1652
USPTO
1911 S. Clark Street, Crystal Mall 1, 10D04, Mail room 10C01
Arlington, VA 22202
(703) 306-0288
delia.ramirez@uspto.gov

THIS PAGE BLANK (USPTO)

=> d his

(FILE 'HOME' ENTERED AT 15:57:10 ON 15 MAY 2002)

INDEX 'ADISALERTS, ADISINSIGHT, ADISNEWS, AGRICOLA, ANABSTR, AQUASCI,
BIOBUSINESS, BIOCOMMERCE, BIOSIS, BIOTECHABS, BIOTECHDS, BIOTECHNO, CABA,
CANCERLIT, CAPLUS, CEABA-VTB, CEN, CIN, CONFSCI, CROPB, CROPU, DDFB,
DDFU, DGENE, DRUGB, DRUGLAUNCH, DRUGMONOG2, ...' ENTERED AT 15:57:19 ON
15 MAY 2002

SEA HUMAN? AND KINAS?

475 FILE ADISALERTS
295 FILE ADISINSIGHT
54 FILE ADISNEWS
297 FILE AGRICOLA
96 FILE ANABSTR
143 FILE AQUASCI
407 FILE BIOBUSINESS
141 FILE BIOCOMMERCE
76512 FILE BIOSIS
1664 FILE BIOTECHABS
1664 FILE BIOTECHDS
32983 FILE BIOTECHNO
1347 FILE CABA
34379 FILE CANCERLIT
46764 FILE CAPLUS
112 FILE CEABA-VTB
33 FILE CEN
77 FILE CIN
569 FILE CONFSCI
27 FILE CROPU
394 FILE DDFB
7268 FILE DDFU
52947 FILE DGENE
394 FILE DRUGB
82 FILE DRUGNL
8737 FILE DRUGU
123 FILE DRUGUPDATES
481 FILE EMBAL
61823 FILE EMBASE
18573 FILE ESBIOBASE
28 FILE FROSTI
29 FILE FSTA
37193 FILE GENBANK
15 FILE HEALSAFE
1105 FILE IFIPAT
5593 FILE JICST-EPLUS
71 FILE KOSMET
13635 FILE LIFESCI
8 FILE MEDICONF
79649 FILE MEDLINE
124 FILE NIOSHTIC
362 FILE NTIS
18 FILE OCEAN
20701 FILE PASCAL
111 FILE PHAR
1 FILE PHIC
157 FILE PHIN
1356 FILE PROMT
49394 FILE SCISEARCH
2 FILE SYNTHLINE
30154 FILE TOXCENTER
19045 FILE USPATFULL
55 FILE USPAT2
1770 FILE WPIDS
1770 FILE WPINDEX
34 FILE IPA
186 FILE NAPRALERT
1843 FILE NLDB

THIS PAGE BLANK (USPTO)

L1 QUE HUMAN? AND KINAS?

FILE 'MEDLINE, BIOSIS, EMBASE, DGENE, SCISEARCH, CAPLUS, GENBANK,
CANCERLIT, BIOTECHNO, TOXCENTER, PASCAL, USPATFULL, ESBIODBASE, LIFESCI,
DRUGU, JICST-EPLUS, NLDB, WPIDS' ENTERED AT 16:00:01 ON 15 MAY 2002

L2 296040 S HUMAN? (S) KINAS?
L3 73739 S L2 (S) (RECOMBIN? OR ISOLAT? OR CLON?)
L4 7879 S L3 (S) (SERIN? OR THREON?)
L5 192 S L3 (S) NEK?
L6 286 S L3 (S) NY?
L7 111 DUP REM L5 (81 DUPLICATES REMOVED)
L8 234 DUP REM L6 (52 DUPLICATES REMOVED)
E WALKER
L9 241 S E3
E TURNER/AU
L10 1510 S E4
E TURNER
L11 64367 S E3
E WALKER/AU
L12 0 S E3
E ZAMBROWICZ/AU
L13 1088 S E4
L14 42 S L3 AND (L12 OR L10 OR L13)
L15 8 S L4 AND (L12 OR L10 OR L13)
L16 0 S L7 AND (L12 OR L10 OR L13)
L17 0 S L8 AND (L12 OR L10 OR L13)
L18 30 DUP REM L14 (12 DUPLICATES REMOVED)

THIS PAGE BLANK (USPTO)

Welcome to STN International! Enter x:x

LOGINID:sssptal652dmr

PASSWORD:

TERMINAL (ENTER 1, 2, 3, OR ?):2

* * * * * Welcome to STN International * * * * *

NEWS 1 Web Page URLs for STN Seminar Schedule - N. America
NEWS 2 Jan 25 BLAST(R) searching in REGISTRY available in STN on the Web
NEWS 3 Jan 29 FSTA has been reloaded and moves to weekly updates
NEWS 4 Feb 01 DKILIT now produced by FIZ Karlsruhe and has a new update frequency
NEWS 5 Feb 19 Access via Tymnet and SprintNet Eliminated Effective 3/31/02
NEWS 6 Mar 08 Gene Names now available in BIOSIS
NEWS 7 Mar 22 TOXLIT no longer available
NEWS 8 Mar 22 TRCTHERMO no longer available
NEWS 9 Mar 28 US Provisional Priorities searched with P in CA/CAPLUS and USPATFULL
NEWS 10 Mar 28 LIPINSKI/CALC added for property searching in REGISTRY
NEWS 11 Apr 02 PAPERCHEM no longer available on STN. Use PAPERCHEM2 instead.
NEWS 12 Apr 08 "Ask CAS" for self-help around the clock
NEWS 13 Apr 09 BEILSTEIN: Reload and Implementation of a New Subject Area
NEWS 14 Apr 09 ZDB will be removed from STN
NEWS 15 Apr 19 US Patent Applications available in IFICDB, IFIPAT, and IFIUDB
NEWS 16 Apr 22 Records from IP.com available in CAPLUS, HCAPLUS, and ZCAPLUS
NEWS 17 Apr 22 BIOSIS Gene Names now available in TOXCENTER
NEWS 18 Apr 22 Federal Research in Progress (FEDRIP) now available

NEWS EXPRESS February 1 CURRENT WINDOWS VERSION IS V6.0d,
CURRENT MACINTOSH VERSION IS V6.0a(ENG) AND V6.0Ja(JP),
AND CURRENT DISCOVER FILE IS DATED 05 FEBRUARY 2002
NEWS HOURS STN Operating Hours Plus Help Desk Availability
NEWS INTER General Internet Information
NEWS LOGIN Welcome Banner and News Items
NEWS PHONE Direct Dial and Telecommunication Network Access to STN
NEWS WWW CAS World Wide Web Site (general information)

Enter NEWS followed by the item number or name to see news on that specific topic.

All use of STN is subject to the provisions of the STN Customer agreement. Please note that this agreement limits use to scientific research. Use for software development or design or implementation of commercial gateways or other similar uses is prohibited and may result in loss of user privileges and other penalties.

* * * * * STN Columbus * * * * *

FILE 'HOME' ENTERED AT 15:57:10 ON 15 MAY 2002

=> index bioscience medicine

FILE 'DRUGMONOG' ACCESS NOT AUTHORIZED

COST IN U.S. DOLLARS

SINCE FILE	TOTAL
ENTRY	SESSION
0.21	0.21

FULL ESTIMATED COST

THIS PAGE BLANK (USPTO)

INDEX 'ADISALERTS, ADISINSIGHT, ADISNEWS, AGRICOLA, ANABSTR, AQUASCI,
BIOBUSINESS, BIOCOMMERCE, BIOSIS, BIOTECHABS, BIOTECHDS, BIOTECHNO, CABA,
CANCERLIT, CAPLUS, CEABA-VTB, CEN, CIN, CONFSCI, CROPB, CROPU, DDFB,
DDFU, DGENE, DRUGB, DRUGLAUNCH, DRUGMONOG2, ...'

ENTERED AT 15:57:19 ON 15 MAY 2002

63 FILES IN THE FILE LIST IN STNINDEX

Enter SET DETAIL ON to see search term postings or to view
search error messages that display as 0* with SET DETAIL OFF.

=> s human? and kinas?

475	FILE ADISALERTS
295	FILE ADISINSIGHT
54	FILE ADISNEWS
297	FILE AGRICOLA
96	FILE ANABSTR
143	FILE AQUASCI
407	FILE BIOBUSINESS
141	FILE BIOCOMMERCE
76512	FILE BIOSIS
1664	FILE BIOTECHABS
1664	FILE BIOTECHDS
32983	FILE BIOTECHNO
1347	FILE CABA
34379	FILE CANCERLIT
46764	FILE CAPLUS
112	FILE CEABA-VTB
33	FILE CEN

17 FILES SEARCHED...

77	FILE CIN
569	FILE CONFSCI
27	FILE CROPU
394	FILE DDFB
7268	FILE DDFU
52947	FILE DGENE
394	FILE DRUGB
82	FILE DRUGNL
8737	FILE DRUGU
123	FILE DRUGUPDATES
481	FILE EMBAL
61823	FILE EMBASE

32 FILES SEARCHED...

18573	FILE ES BIOBASE
28	FILE FROSTI
29	FILE FSTA
37193	FILE GENBANK
15	FILE HEALSAFE
1105	FILE IFIPAT
5593	FILE JICST-EPLUS
71	FILE KOSMET
13635	FILE LIFESCI
8	FILE MEDICONF
79649	FILE MEDLINE
124	FILE NIOSHTIC
362	FILE NTIS
18	FILE OCEAN
20701	FILE PASCAL

49 FILES SEARCHED...

111	FILE PHAR
1	FILE PHIC
157	FILE PHIN

THIS PAGE BLANK (USPTO)

1356	FILE PROMT
49394	FILE SCISEARCH
2	FILE SYNTHLINE
30154	FILE TOXCENTER
19045	FILE USPATFULL
55	FILE USPAT2
1770	FILE WPIDS
1770	FILE WPINDEX
34	FILE IPA
186	FILE NAPRALERT
1843	FILE NLDB

58 FILES HAVE ONE OR MORE ANSWERS, 63 FILES SEARCHED IN STNINDEX

L1 QUE HUMAN? AND KINAS?

=> d rank

F1	79649	MEDLINE
F2	76512	BIOSIS
F3	61823	EMBASE
F4	52947	DGENE
F5	49394	SCISEARCH
F6	46764	CAPLUS
F7	37193	GENBANK
F8	34379	CANCERLIT
F9	32983	BIOTECHNO
F10	30154	TOXCENTER
F11	20701	PASCAL
F12	19045	USPATFULL
F13	18573	ESBIOBASE
F14	13635	LIFESCI
F15	8737	DRUGU
F16	7268	DDFU
F17	5593	JICST-EPLUS
F18	1843	NLDB
F19	1770	WPIDS
F20	1770	WPINDEX
F21	1664	BIOTECHABS
F22	1664	BIOTECHDS
F23	1356	PROMT
F24	1347	CABA
F25	1105	IFIPAT
F26	569	CONFSCI
F27	481	EMBAL
F28	475	ADISALERTS
F29	407	BIOBUSINESS
F30	394	DDFB
F31	394	DRUGB
F32	362	NTIS
F33	297	AGRICOLA
F34	295	ADISINSIGHT
F35	186	NAPRALERT
F36	157	PHIN
F37	143	AQUASCI
F38	141	BIOCOMMERCE
F39	124	NIOSHTIC
F40	123	DRUGUPDATES
F41	112	CEABA-VTB
F42	111	PHAR
F43	96	ANABSTR
F44	82	DRUGNL
F45	77	CIN
F46	71	KOSMET

THIS PAGE BLANK (USPTO)

F47	55	USPAT2
F48	54	ADISNEWS
F49	34	IPA
F50	33	CEN
F51	29	FSTA
F52	28	FROSTI
F53	27	CROPU
F54	18	OCEAN
F55	15	HEALSAFE
F56	8	MEDICONF
F57	2	SYNTHLINE
F58	1	PHIC

=> fil f1-f20

COST IN U.S. DOLLARS

SINCE FILE

TOTAL

ENTRY

SESSION

FULL ESTIMATED COST

2.65

2.86

FILE 'MEDLINE' ENTERED AT 16:00:01 ON 15 MAY 2002

FILE 'BIOSIS' ENTERED AT 16:00:01 ON 15 MAY 2002

COPYRIGHT (C) 2002 BIOLOGICAL ABSTRACTS INC. (R)

FILE 'EMBASE' ENTERED AT 16:00:01 ON 15 MAY 2002

COPYRIGHT (C) 2002 Elsevier Science B.V. All rights reserved.

FILE 'DGENE' ENTERED AT 16:00:01 ON 15 MAY 2002

COPYRIGHT (C) 2002 THOMSON DERWENT

FILE 'SCISEARCH' ENTERED AT 16:00:01 ON 15 MAY 2002

COPYRIGHT (C) 2002 Institute for Scientific Information (ISI) (R)

FILE 'CAPLUS' ENTERED AT 16:00:01 ON 15 MAY 2002

USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.

PLEASE SEE "HELP USAGETERMS" FOR DETAILS.

COPYRIGHT (C) 2002 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'GENBANK' ENTERED AT 16:00:01 ON 15 MAY 2002

FILE 'CANCERLIT' ENTERED AT 16:00:01 ON 15 MAY 2002

FILE 'BIOTECHNO' ENTERED AT 16:00:01 ON 15 MAY 2002

COPYRIGHT (C) 2002 Elsevier Science B.V., Amsterdam. All rights reserved.

FILE 'TOXCENTER' ENTERED AT 16:00:01 ON 15 MAY 2002

COPYRIGHT (C) 2002 ACS

FILE 'PASCAL' ENTERED AT 16:00:01 ON 15 MAY 2002

Any reproduction or dissemination in part or in full,

by means of any process and on any support whatsoever

is prohibited without the prior written agreement of INIST-CNRS.

COPYRIGHT (C) 2002 INIST-CNRS. All rights reserved.

FILE 'USPATFULL' ENTERED AT 16:00:01 ON 15 MAY 2002

CA INDEXING COPYRIGHT (C) 2002 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'ESBIOBASE' ENTERED AT 16:00:01 ON 15 MAY 2002

COPYRIGHT (C) 2002 Elsevier Science B.V., Amsterdam. All rights reserved.

FILE 'LIFESCI' ENTERED AT 16:00:01 ON 15 MAY 2002

COPYRIGHT (C) 2002 Cambridge Scientific Abstracts (CSA)

THIS PAGE BLANK (USPTO)

FILE 'DRUGU' ENTERED AT 16:00:01 ON 15 MAY 2002
COPYRIGHT (C) 2002 THOMSON DERWENT

FILE 'DDFU' ACCESS NOT AUTHORIZED

FILE 'JICST-EPLUS' ENTERED AT 16:00:01 ON 15 MAY 2002
COPYRIGHT (C) 2002 Japan Science and Technology Corporation (JST)

FILE 'NLDB' ENTERED AT 16:00:01 ON 15 MAY 2002
COPYRIGHT (C) 2002 Gale Group. All rights reserved.

FILE 'WPIDS' ENTERED AT 16:00:01 ON 15 MAY 2002
COPYRIGHT (C) 2002 THOMSON DERWENT

FILE 'WPINDEX' ACCESS NOT AUTHORIZED

=> s human? (s) kinas?

4 FILES SEARCHED...

9 FILES SEARCHED...

11 FILES SEARCHED...

L2 296040 HUMAN? (S) KINAS?

=> s l2 (s) (recombin? or isolat? or clon?)

4 FILES SEARCHED...

7 FILES SEARCHED...

12 FILES SEARCHED...

L3 73739 L2 (S) (RECOMBIN? OR ISOLAT? OR CLON?)

<-----User Break----->

u

SEARCH ENDED BY USER

=> s l3 (s) (serin? or threon?)

L4 7879 L3 (S) (SERIN? OR THREON?)

=> s l3 (s) nek?

L5 192 L3 (S) NEK?

=> s l3 (s) ny?

L6 286 L3 (S) NY?

=> dup rem l5

DUPLICATE IS NOT AVAILABLE IN 'DGENE, GENBANK'.

ANSWERS FROM THESE FILES WILL BE CONSIDERED UNIQUE

PROCESSING COMPLETED FOR L5

L7 111 DUP REM L5 (81 DUPLICATES REMOVED)

=> dup rem l6

DUPLICATE IS NOT AVAILABLE IN 'DGENE, GENBANK'.

ANSWERS FROM THESE FILES WILL BE CONSIDERED UNIQUE

PROCESSING COMPLETED FOR L6

L8 234 DUP REM L6 (52 DUPLICATES REMOVED)

=> e walke

E1 12 WALKDOWNS/BI

E2 7 WALKDR4/BI

E3 241 --> WALKE/BI

E4 1 WALKEAKOSKI/BI

E5 1 WALKEAYS/BI

E6 15746 WALKED/BI

E7 2 WALKEDED/BI

E8 1 WALKEDFREQUENTLY/BI

THIS PAGE BLANK (USPTO)

E9	1	WALKEDOFF/BI
E10	1	WALKEDOUT/BI
E11	2	WALKEE/BI
E12	1	WALKEFIELD/BI

=> s e3

L9 241 WALKE/BI

=> e turner/au

'AU' IS NOT A VALID EXPAND FIELD CODE FOR FILE 'NLDB'

E1	1	TURNEN R N/AU
E2	1	TURNENS A P/AU
E3	14	--> TURNER/AU
E4	1510	TURNER A/AU
E5	38	TURNER A A/AU
E6	4	TURNER A ALLAN/AU
E7	164	TURNER A B/AU
E8	249	TURNER A C/AU
E9	223	TURNER A D/AU
E10	3	TURNER A D C/AU
E11	1	TURNER A D M S/AU
E12	8	TURNER A DENISE/AU

The indicated field code is not available for EXPAND in this file. To see a list of valid EXPAND field codes, enter HELP SFIELDS at an arrow prompt (=>).

=> s e4

'AU' IS NOT A VALID FIELD CODE

L10 1510 "TURNER A"/AU

=> e turner

E1	1	TURNEPPE/BI
E2	1	TURNEQ/BI
E3	64367	--> TURNER/BI
E4	1	TURNER1/BI
E5	1	TURNER1852/BI
E6	1	TURNER2000/BI
E7	1	TURNER38/BI
E8	471	TURNERA/BI
E9	327	TURNERACEAE/BI
E10	1	TURNERACEAS/BI
E11	1	TURNERACEES/BI
E12	2	TURNERACEOUS/BI

=> s e3

L11 64367 TURNER/BI

=> e walke/au

'AU' IS NOT A VALID EXPAND FIELD CODE FOR FILE 'NLDB'

E1	1	WALKDHUETE W/AU
E2	1	WALKDORFF H R/AU
E3	0	--> WALKE/AU
E4	3	WALKE A/AU
E5	1	WALKE A E/AU
E6	3	WALKE A L/AU
E7	1	WALKE ALISON B/AU
E8	2	WALKE ALLEN L/AU
E9	50	WALKE B/AU
E10	3	WALKE B H/AU
E11	1	WALKE BERNHARD H/AU
E12	1	WALKE BERTHOLD/AU

The indicated field code is not available for EXPAND in this file. To see a list of valid EXPAND field codes, enter HELP

THIS PAGE BLANK (USPTO)

SFIELDS at an arrow prompt (=>).

=> s e3

'AU' IS NOT A VALID FIELD CODE

L12 0 WALKER/AU

=> e zambrowicz/au

'AU' IS NOT A VALID EXPAND FIELD CODE FOR FILE 'NLDB'

E1	1	ZAMBROWIC B P/AU
E2	1	ZAMBROWICA E B/AU
E3	0 -->	ZAMBROWICZ/AU
E4	1088	ZAMBROWICZ B/AU
E5	4	ZAMBROWICZ B E/AU
E6	93	ZAMBROWICZ B P/AU
E7	2	ZAMBROWICZ BRAIN P/AU
E8	89	ZAMBROWICZ BRIAN/AU
E9	30	ZAMBROWICZ BRIAN P/AU
E10	1	ZAMBROWICZ BRIAN PETER/AU
E11	1	ZAMBROWICZ BRYGIDA E/AU
E12	2	ZAMBROWICZ E/AU

The indicated field code is not available for EXPAND in this file. To see a list of valid EXPAND field codes, enter HELP SFIELDS at an arrow prompt (=>).

=> s e4

'AU' IS NOT A VALID FIELD CODE

L13 1088 "ZAMBROWICZ B"/AU

=> d his

(FILE 'HOME' ENTERED AT 15:57:10 ON 15 MAY 2002)

INDEX 'ADISALERTS, ADISINSIGHT, ADISNEWS, AGRICOLA, ANABSTR, AQUASCI, BIOBUSINESS, BIOCOMMERCE, BIOSIS, BIOTECHABS, BIOTECHDS, BIOTECHNO, CABA, CANCERLIT, CAPLUS, CEABA-VTB, CEN, CIN, CONFSCI, CROPB, CROPU, DDFB, DDFU, DGENE, DRUGB, DRUGLAUNCH, DRUGMONOG2, ...' ENTERED AT 15:57:19 ON 15 MAY 2002

SEA HUMAN? AND KINAS?

475	FILE ADISALERTS
295	FILE ADISINSIGHT
54	FILE ADISNEWS
297	FILE AGRICOLA
96	FILE ANABSTR
143	FILE AQUASCI
407	FILE BIOBUSINESS
141	FILE BIOCOMMERCE
76512	FILE BIOSIS
1664	FILE BIOTECHABS
1664	FILE BIOTECHDS
32983	FILE BIOTECHNO
1347	FILE CABA
34379	FILE CANCERLIT
46764	FILE CAPLUS
112	FILE CEABA-VTB
33	FILE CEN
77	FILE CIN
569	FILE CONFSCI
27	FILE CROPU
394	FILE DDFB
7268	FILE DDFU
52947	FILE DGENE
394	FILE DRUGB

THIS PAGE BLANK (USPTO)

82 FILE DRUGNL
 8737 FILE DRUGU
 123 FILE DRUGUPDATES
 481 FILE EMBAL
 61823 FILE EMBASE
 18573 FILE ESBIODBASE
 28 FILE FROSTI
 29 FILE FSTA
 37193 FILE GENBANK
 15 FILE HEALSAFE
 1105 FILE IFIPAT
 5593 FILE JICST-EPLUS
 71 FILE KOSMET
 13635 FILE LIFESCI
 8 FILE MEDICONF
 79649 FILE MEDLINE
 124 FILE NIOSHTIC
 362 FILE NTIS
 18 FILE OCEAN
 20701 FILE PASCAL
 111 FILE PHAR
 1 FILE PHIC
 157 FILE PHIN
 1356 FILE PROMT
 49394 FILE SCISEARCH
 2 FILE SYNTHLINE
 30154 FILE TOXCENTER
 19045 FILE USPATFULL
 55 FILE USPAT2
 1770 FILE WPIDS
 1770 FILE WPINDEX
 34 FILE IPA
 186 FILE NAPRALERT
 1843 FILE NLDB

L1 QUE HUMAN? AND KINAS?

FILE 'MEDLINE, BIOSIS, EMBASE, DGENE, SCISEARCH, CAPLUS, GENBANK,
 CANCERLIT, BIOTECHNO, TOXCENTER, PASCAL, USPATFULL, ESBIODBASE, LIFESCI,
 DRUGU, JICST-EPLUS, NLDB, WPIDS' ENTERED AT 16:00:01 ON 15 MAY 2002

L2 296040 S HUMAN? (S) KINAS?
 L3 73739 S L2 (S) (RECOMBIN? OR ISOLAT? OR CLON?)
 L4 7879 S L3 (S) (SERIN? OR THREON?)
 L5 192 S L3 (S) NEK?
 L6 286 S L3 (S) NY?
 L7 111 DUP REM L5 (81 DUPLICATES REMOVED)
 L8 234 DUP REM L6 (52 DUPLICATES REMOVED)
 E WALKER
 L9 241 S E3
 E TURNER/AU
 L10 1510 S E4
 E TURNER
 L11 64367 S E3
 E WALKER/AU
 L12 0 S E3
 E ZAMBROWICZ/AU
 L13 1088 S E4

=> s l3 and (l12 or l10 or l13)

15 FILES SEARCHED...

L14 42 L3 AND (L12 OR L10 OR L13)

=> s l4 and (L12 OR L10 OR L13)

THIS PAGE BLANK (USPTO)

L15 8 L4 AND (L12 OR L10 OR L13)

=> s 17 and (L12 OR L10 OR L13)

L16 0 L7 AND (L12 OR L10 OR L13)

=> s 18 and (L12 OR L10 OR L13)

L17 0 L8 AND (L12 OR L10 OR L13)

=> d his

(FILE 'HOME' ENTERED AT 15:57:10 ON 15 MAY 2002)

INDEX 'ADISALERTS, ADISINSIGHT, ADISNEWS, AGRICOLA, ANABSTR, AQUASCI, BIOBUSINESS, BIOCOMMERCE, BIOSIS, BIOTECHABS, BIOTECHDS, BIOTECHNO, CABA, CANCERLIT, CAPLUS, CEABA-VTB, CEN, CIN, CONFSCI, CROPB, CROPU, DDFB, DDFU, DGENE, DRUGB, DRUGLAUNCH, DRUGMONOG2, ...' ENTERED AT 15:57:19 ON 15 MAY 2002

SEA HUMAN? AND KINAS?

475 FILE ADISALERTS
295 FILE ADISINSIGHT
54 FILE ADISNEWS
297 FILE AGRICOLA
96 FILE ANABSTR
143 FILE AQUASCI
407 FILE BIOBUSINESS
141 FILE BIOCOMMERCE
76512 FILE BIOSIS
1664 FILE BIOTECHABS
1664 FILE BIOTECHDS
32983 FILE BIOTECHNO
1347 FILE CABA
34379 FILE CANCERLIT
46764 FILE CAPLUS
112 FILE CEABA-VTB
33 FILE CEN
77 FILE CIN
569 FILE CONFSCI
27 FILE CROPU
394 FILE DDFB
7268 FILE DDFU
52947 FILE DGENE
394 FILE DRUGB
82 FILE DRUGNL
8737 FILE DRUGU
123 FILE DRUGUPDATES
481 FILE EMBAL
61823 FILE EMBASE
18573 FILE ESBIODASE
28 FILE FROSTI
29 FILE FSTA
37193 FILE GENBANK
15 FILE HEALSAFE
1105 FILE IFIPAT
5593 FILE JICST-EPLUS
71 FILE KOSMET
13635 FILE LIFESCI
8 FILE MEDICONF
79649 FILE MEDLINE
124 FILE NIOSHTIC
362 FILE NTIS
18 FILE OCEAN
20701 FILE PASCAL

THIS PAGE BLANK (USPTO)

111 FILE PHAR
 1 FILE PHIC
 157 FILE PHIN
 1356 FILE PROMT
 49394 FILE SCISEARCH
 2 FILE SYNTHLINE
 30154 FILE TOXCENTER
 19045 FILE USPATFULL
 55 FILE USPAT2
 1770 FILE WPIDS
 1770 FILE WPINDEX
 34 FILE IPA
 186 FILE NAPRALERT
 1843 FILE NLDB

L1 QUE HUMAN? AND KINAS?

FILE 'MEDLINE, BIOSIS, EMBASE, DGENE, SCISEARCH, CAPLUS, GENBANK,
 CANCERLIT, BIOTECHNO, TOXCENTER, PASCAL, USPATFULL, ESBIODBASE, LIFESCI,
 DRUGU, JICST-EPLUS, NLDB, WPIDS' ENTERED AT 16:00:01 ON 15 MAY 2002

L2 296040 S HUMAN? (S) KINAS?
 L3 73739 S L2 (S) (RECOMBIN? OR ISOLAT? OR CLON?)
 L4 7879 S L3 (S) (SERIN? OR THREON?)
 L5 192 S L3 (S) NEK?
 L6 286 S L3 (S) NY?
 L7 111 DUP REM L5 (81 DUPLICATES REMOVED)
 L8 234 DUP REM L6 (52 DUPLICATES REMOVED)
 E WALKER
 L9 241 S E3
 E TURNER/AU
 L10 1510 S E4
 E TURNER
 L11 64367 S E3
 E WALKER/AU
 L12 0 S E3
 E ZAMBROWICZ/AU
 L13 1088 S E4
 L14 42 S L3 AND (L12 OR L10 OR L13)
 L15 8 S L4 AND (L12 OR L10 OR L13)
 L16 0 S L7 AND (L12 OR L10 OR L13)
 L17 0 S L8 AND (L12 OR L10 OR L13)

=> dup rem l14

DUPLICATE IS NOT AVAILABLE IN 'DGENE, GENBANK'.

ANSWERS FROM THESE FILES WILL BE CONSIDERED UNIQUE

PROCESSING COMPLETED FOR L14

L18 30 DUP REM L14 (12 DUPLICATES REMOVED)

=> d ti l18 1-30

L18 ANSWER 1 OF 30 WPIDS (C) 2002 THOMSON DERWENT

TI Novel **isolated human kinase** polynucleotide
 that shares structural similarity with animal **kinases** including
 calcium/calmodulin-dependent protein **kinases** and
 serine/threonine protein **kinases**, useful in therapeutics.

L18 ANSWER 2 OF 30 WPIDS (C) 2002 THOMSON DERWENT

TI Novel **isolated human kinase** polynucleotide
 useful for screening for drugs effective in treatment of symptomatic or
 phenotypic manifestations of perturbing normal function of novel
human kinase protein in the body.

L18 ANSWER 3 OF 30 WPIDS (C) 2002 THOMSON DERWENT

THIS PAGE BLANK (USPTO)

TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and for
screening for compounds useful for treating mental, biological or medical
diseases.

L18 ANSWER 4 OF 30 WPIDS (C) 2002 THOMSON DERWENT
TI New substituted 1,4-Dihydroindeno(1,2-C)Pyrazoles useful in the treatment
of e.g. cancer, arthritis and atherosclerosis are Tyrosine Kinase
Inhibitors.

L18 ANSWER 5 OF 30 WPIDS (C) 2002 THOMSON DERWENT
TI Constructing normalized eukaryotic cDNA libraries, useful for isolating
and identifying new cDNA.

L18 ANSWER 6 OF 30 MEDLINE DUPLICATE 1
TI Characterization of a multidrug resistant human erythroleukemia cell line
(K562) exhibiting spontaneous resistance to 1-beta-D-
arabinofuranosylcytosine.

L18 ANSWER 7 OF 30 MEDLINE DUPLICATE 2
TI In vitro effects of bryostatin 1 on the metabolism and cytotoxicity of
1-beta-D-arabinofuranosylcytosine in human leukemia cells.

L18 ANSWER 8 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI An isolated nucleic acid molecule encoding a novel human protein useful
as therapeutics and to screen libraries isolate clones and prepare
cloning and sequencing templates -

L18 ANSWER 9 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI Novel **isolated human kinase** polynucleotide
that shares structural similarity with animal **kinases** including
calcium/calmodulin-dependent protein **kinases** and
serine/threonine protein **kinases**, useful in therapeutics -

L18 ANSWER 10 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI Novel **isolated human kinase** polynucleotide
that shares structural similarity with animal **kinases** including
calcium/calmodulin-dependent protein **kinases** and
serine/threonine protein **kinases**, useful in therapeutics -

L18 ANSWER 11 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI Novel **isolated human kinase** polynucleotide
useful for screening for drugs effective in treatment of symptomatic or
phenotypic manifestations of perturbing normal function of novel
human kinase protein in the body -

L18 ANSWER 12 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or
medical diseases -

L18 ANSWER 13 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or
medical diseases -

L18 ANSWER 14 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or
medical diseases -

THIS PAGE BLANK (USPTO)

L18 ANSWER 15 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or
medical diseases -

L18 ANSWER 16 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or
medical diseases -

L18 ANSWER 17 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or
medical diseases -

L18 ANSWER 18 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI An isolated nucleic acid molecule encoding a novel human protein useful
as therapeutics and to screen libraries isolate clones and prepare
cloning and sequencing templates -

L18 ANSWER 19 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI An isolated nucleic acid molecule encoding a novel human protein useful
as therapeutics and to screen libraries isolate clones and prepare
cloning and sequencing templates -

L18 ANSWER 20 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI Novel **isolated human kinase** polynucleotide
that shares structural similarity with animal **kinases** including
calcium/calmodulin-dependent protein **kinases** and
serine/threonine protein **kinases**, useful in therapeutics -

L18 ANSWER 21 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI Novel **isolated human kinase** polynucleotide
that shares structural similarity with animal **kinases** including
calcium/calmodulin-dependent protein **kinases** and
serine/threonine protein **kinases**, useful in therapeutics -

L18 ANSWER 22 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI Novel **isolated human kinase** polynucleotide
that shares structural similarity with animal **kinases** including
calcium/calmodulin-dependent protein **kinases** and
serine/threonine protein **kinases**, useful in therapeutics -

L18 ANSWER 23 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI Novel **isolated human kinase** polynucleotide
useful for screening for drugs effective in treatment of symptomatic or
phenotypic manifestations of perturbing normal function of novel
human kinase protein in the body -

L18 ANSWER 24 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or
medical diseases -

L18 ANSWER 25 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or

THIS PAGE BLANK (USPTO)

medical diseases -

L18 ANSWER 26 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or
medical diseases -

L18 ANSWER 27 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or
medical diseases -

L18 ANSWER 28 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or
medical diseases -

L18 ANSWER 29 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or
medical diseases -

L18 ANSWER 30 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or
medical diseases -

=> d 1

L18 ANSWER 1 OF 30 WPIDS (C) 2002 THOMSON DERWENT
AN 2001-381667 [40] WPIDS
DNC C2001-116955
TI Novel **isolated human kinase** polynucleotide
that shares structural similarity with animal **kinases** including
calcium/calmodulin-dependent protein **kinases** and
serine/threonine protein **kinases**, useful in therapeutics.
DC B04 D16
IN ABUIN, A; DONOHO, G; FRIEDRICH, G; SANDS, A T; SCOVILLE, J; TURNER, C A;
ZAMBROWICZ, B
PA (LEXI-N) LEXICON GENETICS INC
CYC 93
PI WO 2001042435 A2 20010614 (200140)* EN 31p C12N009-00
RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ
NL OA PT SD SE SL SZ TR TZ UG ZW
W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CR CU CZ DE DK DM
DZ EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC
LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE
SG SI SK SL TJ TM TR TT TZ UA UG UZ VN YU ZA ZW
AU 2001025767 A 20010618 (200161) C12N009-00
ADT WO 2001042435 A2 WO 2000-US33240 20001207; AU 2001025767 A AU 2001-25767
20001207
FDT AU 2001025767 A Based on WO 200142435
PRAI US 1999-169428P 19991207
IC ICM C12N009-00

=> d 4-9

THIS PAGE BLANK (USPTO)

L18 ANSWER 4 OF 30 WPIDS (C) 2002 THOMSON DERWENT
 AN 2000-647337 [62] WPIDS
 DNC C2000-195868
 TI New substituted 1,4-Dihydroindeno(1,2-C)Pyrazoles useful in the treatment of e.g. cancer, arthritis and atherosclerosis are Tyrosine Kinase Inhibitors.
 DC B02
 IN ARNOLD, L; DOYLE, K; RAFFERTY, P; STEELE, R; **TURNER, A**; WILKINS, D; ARNOLD, L D; DOYLE, K J; STEELE, R W; WILKINS, D J
 PA (BADI) BASF AG; (KNOL) KNOLL GMBH
 CYC 43
 PI WO 2000059901 A1 20001012 (200062)* EN 78p C07D401-12
 RW: AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE
 W: AU BG BR CA CN CZ HR HU ID IL IN JP KR MX NO NZ PL RU SG SK TR UA
 US ZA
 AU 2000040378 A 20001023 (200107) C07D401-12
 US 6297238 B1 20011002 (200160) A61K031-4162
 EP 1165544 A1 20020102 (200209) EN C07D401-12
 R: AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
 NO 2001004864 A 20011205 (200210) C07D000-00
 CZ 2001003580 A3 20020313 (200223) C07D495-04
 ADT WO 2000059901 A1 WO 2000-US8192 20000328; AU 2000040378 A AU 2000-40378 20000328; US 6297238 B1 Provisional US 1999-127963P 19990406, CIP of US 2000-541336 20000403, US 2000-689943 20001012; EP 1165544 A1 EP 2000-919742 20000328, WO 2000-US8192 20000328; NO 2001004864 A WO 2000-US8192 20000328, NO 2001-4864 20011005; CZ 2001003580 A3 WO 2000-US8192 20000328, CZ 2001-3580 20000328
 FDT AU 2000040378 A Based on WO 200059901; EP 1165544 A1 Based on WO 200059901; CZ 2001003580 A3 Based on WO 200059901
 PRAI US 1999-127963P 19990406; US 2000-541336 20000403; US 2000-689943 20001012
 IC ICM A61K031-4162; C07D000-00; C07D401-12; C07D495-04
 ICS A61K031-416; A61P001-04; A61P009-10; A61P031-04; A61P035-00; A61P043-00; C07D231-54; C07D403-12

L18 ANSWER 5 OF 30 WPIDS (C) 2002 THOMSON DERWENT
 AN 2000-224329 [19] WPIDS
 DNC C2000-068498
 TI Constructing normalized eukaryotic cDNA libraries, useful for isolating and identifying new cDNA.
 DC B04 D16
 IN FRIEDRICH, G; NEHLS, M; RULEY, H E; SANDS, A T; WATTLER, S;
ZAMBROWICZ, B
 PA (LEXI-N) LEXICON GENETICS INC
 CYC 87
 PI WO 2000009681 A2 20000224 (200019)* EN 47p C12N015-10
 RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW NL
 OA PT SD SE SL SZ UG ZW
 W: AE AL AM AT AU AZ BA BB BG BR BY CA CH CN CU CZ DE DK EE ES FI GB
 GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU
 LV MD MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR
 TT UA UG UZ VN YU ZA ZW
 AU 9954703 A 20000306 (200030) C12N015-10
 US 6218123 B1 20010417 (200123) C12Q001-68
 EP 1105471 A2 20010613 (200134) EN C12N015-10
 R: AL AT BE CH CY DE DK ES FI FR GB GR IE IT LI LT LU LV MC MK NL PT
 RO SE SI
 ADT WO 2000009681 A2 WO 1999-US17945 19990810; AU 9954703 A AU 1999-54703 19990810; US 6218123 B1 Provisional US 1998-95989P 19980810, US 1999-371257 19990810; EP 1105471 A2 EP 1999-940956 19990810, WO 1999-US17945 19990810
 FDT AU 9954703 A Based on WO 200009681; EP 1105471 A2 Based on WO 200009681

THIS PAGE BLANK (USPTO)

PRAI US 1998-95989P 19980810; US 1999-371257 19990810
IC ICM C12N015-10; C12Q001-68
ICS C12N015-11; C12N015-62; C12N015-63

L18 ANSWER 6 OF 30 MEDLINE DUPLICATE 1
AN 95287645 MEDLINE
DN 95287645 PubMed ID: 7769843
TI Characterization of a multidrug resistant human erythroleukemia cell line (K562) exhibiting spontaneous resistance to 1-beta-D-arabinofuranosylcytosine.
AU Grant S; **Turner A**; Nelms P; Yanovich S
CS Department of Medicine, Medical College of Virginia, Richmond 23298, USA.
NC 1R01 CA63753 (NCI)
CA-16059 (NCI)
SO LEUKEMIA, (1995 May) 9 (5) 808-14.
Journal code: LEU; 8704895. ISSN: 0887-6924.
CY ENGLAND: United Kingdom
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 199507
ED Entered STN: 19950713
Last Updated on STN: 19970203
Entered Medline: 19950706

L18 ANSWER 7 OF 30 MEDLINE DUPLICATE 2
AN 91328750 MEDLINE
DN 91328750 PubMed ID: 1867641
TI In vitro effects of bryostatin 1 on the metabolism and cytotoxicity of 1-beta-D-arabinofuranosylcytosine in human leukemia cells.
AU Grant S; Boise L; Westin E; Howe C; Pettit G R; **Turner A**; McCrady C
CS Division of Hematology/Oncology, Medical College of Virginia, Richmond 23298.
NC AICR88B36 (NIAID)
CA04875 (NCI)
R01 CA35601 (NCI)
SO BIOCHEMICAL PHARMACOLOGY, (1991 Jul 25) 42 (4) 853-67.
Journal code: 9Z4; 0101032. ISSN: 0006-2952.
CY ENGLAND: United Kingdom
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 199109
ED Entered STN: 19910929
Last Updated on STN: 19970203
Entered Medline: 19910909

L18 ANSWER 8 OF 30 DGENE (C) 2002 THOMSON DERWENT
AN AAU04636 Protein DGENE
TI An isolated nucleic acid molecule encoding a novel human protein useful as therapeutics and to screen libraries isolate clones and prepare cloning and sequencing templates -
IN Donoho G; Hilbun E; Turner C A; Friedrich G; **Zambrowicz B**; Sands A T
PA (LEXI-N) LEXICON GENETICS INC.
PI WO 2001053493 A2 20010726 33p
AI WO 2001-US2120 20010118
PRAI US 2000-176690 20000118
DT Patent
LA English
OS 2001-442260 [47]

THIS PAGE BLANK (USPTO)

L18 ANSWER 9 OF 30 DGENE (C) 2002 THOMSON DERWENT
 AN AAB84360 Protein DGENE
 TI Novel **isolated human kinase** polynucleotide
 that shares structural similarity with animal **kinases** including
 calcium/calmodulin-dependent protein **kinases** and
 serine/threonine protein **kinases**, useful in therapeutics -
 IN Donoho G; Scoville J; Turner C A; Friedrich G; **Zambrowicz B**;
 Abuin A; Sands A T
 PA (LEXI-N) LEXICON GENETICS INC.
 PI WO 2001042435 A2 20010614 32p
 AI WO 2000-US33240 20001207
 PRAI US 1999-169428 19991207
 DT Patent
 LA English
 OS 2001-381667 [40]

=> d his

(FILE 'HOME' ENTERED AT 15:57:10 ON 15 MAY 2002)

INDEX 'ADISALERTS, ADISINSIGHT, ADISNEWS, AGRICOLA, ANABSTR, AQUASCI,
 BIOBUSINESS, BIOCOMMERCE, BIOSIS, BIOTECHABS, BIOTECHDS, BIOTECHNO, CABA,
 CANCERLIT, CAPLUS, CEABA-VTB, CEN, CIN, CONFSCI, CROPB, CROPU, DDFB,
 DDFU, DGENE, DRUGB, DRUGLAUNCH, DRUGMONOG2, ...' ENTERED AT 15:57:19 ON
 15 MAY 2002

SEA HUMAN? AND KINAS?

 475 FILE ADISALERTS
 295 FILE ADISINSIGHT
 54 FILE ADISNEWS
 297 FILE AGRICOLA
 96 FILE ANABSTR
 143 FILE AQUASCI
 407 FILE BIOBUSINESS
 141 FILE BIOCOMMERCE
 76512 FILE BIOSIS
 1664 FILE BIOTECHABS
 1664 FILE BIOTECHDS
 32983 FILE BIOTECHNO
 1347 FILE CABA
 34379 FILE CANCERLIT
 46764 FILE CAPLUS
 112 FILE CEABA-VTB
 33 FILE CEN
 77 FILE CIN
 569 FILE CONFSCI
 27 FILE CROPU
 394 FILE DDFB
 7268 FILE DDFU
 52947 FILE DGENE
 394 FILE DRUGB
 82 FILE DRUGNL
 8737 FILE DRUGU
 123 FILE DRUGUPDATES
 481 FILE EMBAL
 61823 FILE EMBASE
 18573 FILE ESBIODBASE
 28 FILE FROSTI
 29 FILE FSTA
 37193 FILE GENBANK
 15 FILE HEALSAFE
 1105 FILE IFIPAT

THIS PAGE BLANK (USPTO)

5593 FILE JICST-EPLUS
 71 FILE KOSMET
 13635 FILE LIFESCI
 8 FILE MEDICONF
 79649 FILE MEDLINE
 124 FILE NIOSHTIC
 362 FILE NTIS
 18 FILE OCEAN
 20701 FILE PASCAL
 111 FILE PHAR
 1 FILE PHIC
 157 FILE PHIN
 1356 FILE PROMT
 49394 FILE SCISEARCH
 2 FILE SYNTHLINE
 30154 FILE TOXCENTER
 19045 FILE USPATFULL
 55 FILE USPAT2
 1770 FILE WPIDS
 1770 FILE WPINDEX
 34 FILE IPA
 186 FILE NAPRALERT
 1843 FILE NLDB

L1 QUE HUMAN? AND KINAS?

FILE 'MEDLINE, BIOSIS, EMBASE, DGENE, SCISEARCH, CAPLUS, GENBANK,
 CANCERLIT, BIOTECHNO, TOXCENTER, PASCAL, USPATFULL, ESBIODBASE, LIFESCI,
 DRUGU, JICST-EPLUS, NLDB, WPIDS' ENTERED AT 16:00:01 ON 15 MAY 2002

L2 296040 S HUMAN? (S) KINAS?
 L3 73739 S L2 (S) (RECOMBIN? OR ISOLAT? OR CLON?)
 L4 7879 S L3 (S) (SERIN? OR THREON?)
 L5 192 S L3 (S) NEK?
 L6 286 S L3 (S) NY?
 L7 111 DUP REM L5 (81 DUPLICATES REMOVED)
 L8 234 DUP REM L6 (52 DUPLICATES REMOVED)
 E WALKER
 L9 241 S E3
 E TURNER/AU
 L10 1510 S E4
 E TURNER
 L11 64367 S E3
 E WALKER/AU
 L12 0 S E3
 E ZAMBROWICZ/AU
 L13 1088 S E4
 L14 42 S L3 AND (L12 OR L10 OR L13)
 L15 8 S L4 AND (L12 OR L10 OR L13)
 L16 0 S L7 AND (L12 OR L10 OR L13)
 L17 0 S L8 AND (L12 OR L10 OR L13)
 L18 30 DUP REM L14 (12 DUPLICATES REMOVED)

=> log h

COST IN U.S. DOLLARS

SINCE FILE	TOTAL
ENTRY	SESSION
97.00	99.86

FULL ESTIMATED COST

SESSION WILL BE HELD FOR 60 MINUTES
 STN INTERNATIONAL SESSION SUSPENDED AT 16:25:25 ON 15 MAY 2002

THIS PAGE BLANK (USPTO)

**PALM INTRANET**

Day : Wednesday
Date: 5/15/2002
Time: 16:03:10

Inventor Name Search

Enter the **first few letters** of the Inventor's Last Name.
Additionally, enter the **first few letters** of the Inventor's First name.

Last Name**First Name**

(To go back use Back button on your browser toolbar.)

Back to [PALM](#) | [ASSIGNMENT](#) | [OASIS](#) | [Home page](#)

THIS PAGE BLANK (USPTO)

**PALM INTRANET**

Day : Wednesday
Date: 5/15/2002
Time: 16:03:10

Inventor Name Search

Enter the **first few letters** of the Inventor's Last Name.
Additionally, enter the **first few letters** of the Inventor's First name.

Last Name**First Name**

(To go back use Back button on your browser toolbar.)

Back to [PALM](#) | [ASSIGNMENT](#) | [OASIS](#) | [Home page](#)

THIS PAGE BLANK (USPTO)

**PALM INTRANET**

Day : Wednesday
Date: 5/15/2002
Time: 16:03:10

Inventor Name Search

Enter the **first few letters** of the Inventor's Last Name.
Additionally, enter the **first few letters** of the Inventor's First name.

Last Name**First Name**

(To go back use Back button on your browser toolbar.)

Back to [PALM](#) | [ASSIGNMENT](#) | [OASIS](#) | [Home page](#)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 15, 2002, 07:55:13 ; Search time 93.14 Seconds
(without alignments)
2254.840 Million cell updates/sec

Title: US-09-783-320-4
Perfect score: 62.3
Sequence: 1 MEKTVRLQKIGESFGKAIL.....YAKILHLVWADGAYCEDNDE 1214

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL.19:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.organelle:*
10: sp.phage:*
11: sp.plant:*
12: sp.podent:*
13: sp.yarnus:*
14: sp.vertebrate:*
15: sp.unclassified:*
16: sp.virus:*
17: sp.bacteriap:*
18: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6201	99.3	1265	4	Q96PY6
2	3945	63.2	771	4	Q9Y594
3	1953	31.3	375	4	Q9H6P7
4	1387	22.2	320	4	Q96S52
5	836	13.4	744	11	Q35673
6	836	13.4	792	11	Q9R1J1
7	836	13.4	792	11	Q921J2
8	827	13.2	509	11	Q99K72
9	695	11.1	555	5	Q9N9C3
10	622.5	10.0	698	11	Q912R4
11	621.5	9.9	606	10	Q9CAU7
12	617	9.9	697	13	Q90XC2
13	614	9.8	291	11	Q9D6B5
14	609	9.8	943	10	Q94C05
15	596.5	9.6	609	10	Q947T1
16	593	9.5	312	4	Q9Y6S4

17	590.5	9.5	579	5	Q95XQ3	Q95XQ3 caenorhabdi
18	584	9.4	841	5	Q9VC32	Q9VC32 drosophila
19	566.5	9.1	416	10	Q9LR35	Q9LR35 arabidopsis
20	559.5	9.0	561	5	Q76134	Q76134 tetrahymena
21	557	8.9	941	10	Q9LXP3	Q9LXP3 arabidopsis
22	555.5	8.9	443	11	Q91218	Q91218 mus musculus
23	555	8.9	442	13	Q9W622	Q9W622 xenopus lae
24	548	8.8	384	4	Q960N9	Q960N9 homo sapien
25	546.5	8.8	443	11	Q921N9	Q921N9 mus musculus
26	544.5	8.7	389	13	Q9W623	Q9W623 xenopus lae
27	540.5	8.7	357	5	Q01775	Q01775 caenorhabdi
28	520	8.3	735	5	Q9W41	Q9W41 rattus norv
29	519	8.3	723	3	Q9W41	Q9W41 drosophila
30	514.5	8.0	1232	11	Q9K389	Q9K389 schizosacch
31	500	7.9	1257	4	Q9K389	Q9K389 mus musculus
32	494.5	7.9	1352	4	Q9K389	Q9K389 mus musculus
33	494	7.9	1352	4	Q9K389	Q9K389 mus musculus
34	493.5	7.9	1202	11	Q9W41	Q9W41 mus musculus
35	490.5	7.9	1305	4	Q9K389	Q9K389 mus musculus
36	490	7.8	1268	4	Q9K389	Q9K389 mus musculus
37	490	7.8	1333	4	Q9K389	Q9K389 mus musculus
38	490	7.8	1333	4	Q9K389	Q9K389 mus musculus
39	487.5	7.8	1235	4	Q9K389	Q9K389 mus musculus
40	487.5	7.8	1235	4	Q9K389	Q9K389 mus musculus
41	486	7.8	1276	4	Q9K389	Q9K389 mus musculus
42	483.5	7.7	1331	4	Q9K389	Q9K389 mus musculus
43	482.5	7.7	1204	4	Q9K389	Q9K389 mus musculus
44	479.5	7.7	306	4	Q9ULX2	Q9ULX2 homo sapien
45	479.5	7.7	313	4	Q96E83	Q96E83 homo sapien

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	1265 AA.
Q96PY6	1	Q96PY6		
ID	Q96PY6	Q96PY6		
AC	Q96PY6	Q96PY6		
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	KIAI1901 PROTEIN (FRAGMENT).			
GN	KIAI1901.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=BRAIN;			
RX	MEDLINE=21456161; PubMed=11572484;			
RA	Nagase T., Kikuno R., Ohara O.;			
RT	"Prediction of the coding sequences of unidentified human genes. XXI.			
RT	"The complete sequences of 60 new cDNA clones from brain which code for			
RT	large proteins."			
RL	DNA Res. 8:179-187(2001).			
DR	EMBL; AB067488; BAB67794.1; -			
FT	NON_TER			
SQ	SEQUENCE	1265 AA; 143754 MW; 31DDF2DFEAE39C89 CRC64;		

QY	1	MEKTVRLQKIGESFGKAILVSTEDGRQYVIRKEINISRMSSKERESRRVAVLANMKH 60	99.3%; Score 6201; DB 4; Length 1265;
DB	8	MEKTVRLQKIGESFGKAILVSTEDGRQYVIRKEINISRMSSKERESRRVAVLANMKH 67	Best local Similarity 96.3%; Pred. No. 6.5e-314;
QY	61	PNIVQYRSPFENSGSLIVMDYCEGDLFRINAGKGVLFEDQITLDFVQICLAKHVH 120	Mismatches 1; Indels 44; Gaps 1;
DB	68	PNIVQYRSPFENSGSLIVMDYCEGDLFRINAGKGVLFEDQITLDFVQICLAKHVH 127	

Wed May 15 14:14:06 2002

us-09-783-320-4.rspt

Page 2

```
121 DRKILRDIKSONIFLTCKDVTGOLGDFIARVNSTVELARTCTGTPYLSPECNKPY 180
128 DRKILRDIKSONIFLTCKDVTGOLGDFIARVNSTVELARTCTGTPYLSPECNKPY 187
181 NKSQDIAWALGCVLYELCTAKHAFEGASMKNVLYKTSSFPVLYHSYDLRSVQLFK 240
188 NKSQDIAWALGCVLYELCTAKHAFEGASMKNVLYKTSSFPVLYHSYDLRSVQLFK 247
241 RNRDRPVSNSILEKFIKRIEKLSPOLLAEEFCLTKFSKFSQOPAPRAPSGONSI 300
248 RNRDRPVSNSILEKFIKRIEKLSPOLLAEEFCLTKFSKFSQOPAPRAPSGONSI 307
301 SVMPAOKITRPAKYGIPLAYKKGDKLHEKKPLQKHQAQHTPEKRVNTGEBRRRISE 360
308 SVMPAOKITRPAKYGIPLAYKKGDKLHEKKPLQKHQAQHTPEKRVNTGEBRRRISE 367
361 EAARKRLEFIEKEKKOKDOITSLMKAQOMKROEKERLERINRAEQGMNVLSAGSGE 420
368 EAARKRLEFIEKEKKOKDOITSLMKAQOMKROEKERLERINRAEQGMNVLSAGSGE 427
421 VKAPFISGSGTAPSSFSRGQYEHYHAFIDMOQOARADNEAKMKREIYRGILPE--- 476
428 VKAPFISGSGTAPSSFSRGQYEHYHAFIDMOQOARADNEAKMKREIYRGILPEKGI 487
477 -----ROKQOLAVERAQVEEFLOR 496
488 PGVRPGFYGAAGHHFPPADDIRTKLRKRAVSKQANANROKQOLAVERAQVEEFLOR 547
497 KRAMONKARABGHMYLRLRQILQNFNROQIKAKLGEKKANHSQEGSEADAM 556
548 KRAMONKARABGHMYLRLRQILQNFNROQIKAKLGEKKANHSQEGSEADAM 607
557 RRRKIESLAHANARAVALKEQLERKREKREKREKREKREKREKREKREKREKREK 616
608 RRRKIESLAHANARAVALKEQLERKREKREKREKREKREKREKREKREKREKREK 667
617 TGSPSKQOMRSVIVTSALKLEVGVDSLTDTRETSEMOKTNNATISSKREILRLNENL 676
668 TGSPSKQOMRSVIVTSALKLEVGVDSLTDTRETSEMOKTNNATISSKREILRLNENL 727
677 KAOEDKGMONLSDTFEIVNHEDAEHEKREKREKREKREKREKREKREKREKREKREK 736
728 KAOEDKGMONLSDTFEIVNHEDAEHEKREKREKREKREKREKREKREKREKREKREK 787
737 TTRRHVGEYIKLGPNGSPRRAMGSPDVSYLKILGEAELOLOTELENTTIRSEISPEG 796
788 TTRRHVGEYIKLGPNGSPRRAMGSPDVSYLKILGEAELOLOTELENTTIRSEISPEG 847
797 EKRPRLITGEKKVOCISHEINSAIVDSPEVETKSPSEASPOMSKLEGNLEPDDLET 856
848 EKRPRLITGEKKVOCISHEINSAIVDSPEVETKSPSEASPOMSKLEGNLEPDDLET 907
857 EIIQEPSTGNKDESLPCTITDVAISEKETKETOSADRITTOENEVSEDOGVSTVDOLSD 916
908 EIIQEPSTGNKDESLPCTITDVAISEKETKETOSADRITTOENEVSEDOGVSTVDOLSD 967
917 IHIEPSTGNKDESLPCTITDVAISEKETKETOSADRITTOENEVSEDOGVSTVDOLSD 976
968 IHIEPSTGNKDESLPCTITDVAISEKETKETOSADRITTOENEVSEDOGVSTVDOLSD 1027
977 LPRKNNKNSLILGLSTGLFDANNPKMLRTCSLPDLSKLFRTLMQVPGVGDVRODNLEID 1036
1028 LPRKNNKNSLILGLSTGLFDANNPKMLRTCSLPDLSKLFRTLMQVPGVGDVRODNLEID 1087
1037 EIRDENIKGSPDSEDIIVEEDTDLOLOASMEOLLEBOGGESESEESVYLNKSNYER 1096
1088 EIRDENIKGSPDSEDIIVEEDTDLOLOASMEOLLEBOGGESESEESVYLNKSNYER 1147
1097 TANGTVADEDDNPSSESLNNEHMSDNDGFIASECDSVFNHLELRLHLQEBKGF 1156
1148 TANGTVADEDDNPSSESLNNEHMSDNDGFIASECDSVFNHLELRLHLQEBKGF 1207
1157 KFEFVEYERIKAIHEDEDENIETCSKIVONITLGNHQLYARILHLVMAAGVQOEDNDE 1214
```

```
1208 KFEFVEYERIKAIHEDEDENIETCSKIVONITLGNHQLYARILHLVMAAGVQOEDNDE 1265

RESULT 2
ID 09Y594 PRELIMINARY; PRT: 771 AA.
AC 09Y594: 1999 (TREMBLER. 12, Created)
DT 01-NOV-1999 (TREMBLER. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLER. 19, Last annotation update)
DE NY-REN-55 ANTIGEN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
OX 11
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=99438124; Pubmed=10508479;
RA Scanlan M.J., Gordon J.D., Williamson B., Stockert E., Bander N.H.,
RA Jorgensen V., Gure A.O., Jager D., Jager E., Knuth A., Chen Y.-T.,
RA Old L.J.;
RT "Antigens recognized by autologous antibody in patients with renal-
cell carcinoma";
RL Int. J. Cancer 83:456-464(1999).
EMBL: AF15113; AAD42879.1; -.
FT NONTER 1
SQ SEQUENCE 771 AA; 87539 MW; 95D76131F432F6CF CRC64;

Query Match 63.28; Score 3945; DB 4; Length 771;
Best Local Similarity 99.68; Pred. No. 5e-197;
Matches 768; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

444 EHYHAFIDMOQOARADNEAKMKREIYRGILPEROKQOLAVERAQVEEFLORREAMON 503
1 EHYHAFIDMOQOARADNEAKMKREIYRGILPEROKQOLAVERAQVEEFLORREAMON 60
504 KARAGGHVYLYLARLQILQNFNROQIKAKLGEKKANHSQEGSEADAMRRKIES 563
61 KARAGGHVYLYLARLQILQNFNROQIKAKLGEKKANHSQEGSEADAMRRKIES 120
564 LKAHANARAVALKEQLERKREKREKREKREKREKREKREKREKREKREKREKREK 623
121 LKAHANARAVALKEQLERKREKREKREKREKREKREKREKREKREKREKREKREK 180
624 OOMRSVIVTSALKLEVGVDSLTDTRETSEMOKTNNATISSKREILRLNENLKAODEK 683
181 OOMRSVIVTSALKLEVGVDSLTDTRETSEMOKTNNATISSKREILRLNENLKAODEK 240
684 GMONLSDTFEIVNHEDAEHEKREKREKREKREKREKREKREKREKREKREKREKRE 743
241 GMONLSDTFEIVNHEDAEHEKREKREKREKREKREKREKREKREKREKREKREKRE 300
744 GEVILGPNGSPRRAMGSPDVSYLKILGEAELOLOTELENTTIRSEISPEGKPKLI 803
301 GEVILGPNGSPRRAMGSPDVSYLKILGEAELOLOTELENTTIRSEISPEGKPKLI 360
804 TEKKVOCISHEINSAIVDSPEVETKSPSEASPOMSKLEGNLEPDDLETIELOEPS 863
361 TEKKVOCISHEINSAIVDSPEVETKSPSEASPOMSKLEGNLEPDDLETIELOEPS 420
864 GNRKRESLPCTITDVAISEKETKETOSADRITTOENEVSEDOGVSTVDOLSDIHIEP 923
421 GNRKRESLPCTITDVAISEKETKETOSADRITTOENEVSEDOGVSTVDOLSDIHIEP 480
924 NDSQSKCDVKSVOPEPFHKKVYHSEHLVPOVOSVOCSPSEASVFNHLELRLHLQEBK 983
481 NDSQSKCDVKSVOPEPFHKKVYHSEHLVPOVOSVOCSPSEASVFNHLELRLHLQEBK 540
984 KNSLLIGLSTGLFDANNPKMLRTCSLPDLSKLFRTLMQVPGVGDVRODNLEIDETDENI 1043
541 KNSLLIGLSTGLFDANNPKMLRTCSLPDLSKLFRTLMQVPGVGDVRODNLEIDETDENI 600
```

QY 1044 KEGPDSSEDIIVFEETDTDLOELQASMEQLLRQEPGESESEESVYLNKNSDVEPTANGTDV 1103
 |||||||
 DB 601 KEGPDSSEDIIVFEETDTDLOELQASMEQLLRQEPGESESEESVYLNKNSDVEPTANGTDV 660
 |||||||
 QY 1104 ADEPDNPSSESSALNEMHSDNSNGEJASSECDSVFNHLEELRLHLEQEMGEKFEFFVE 1163
 |||||||
 DB 661 ADEPDNPSSESSALNEMHSDNSNGEJASSECDSVFNHLEELRLHLEQEMGEKFEFFVE 720
 |||||||
 QY 1164 KIRKIHEDDENIEICISKYONITLGNHGHOLYAKITLHVMADGAYOEONDE 1214
 |||||||
 DB 721 KIRKIHEDDENIEICISKYONITLGNHGHOLYAKITLHVMADGAYOEONDE 771
 |||||||
 RESULT 3
 Q9H6P7 PRELIMINARY: PRT: 375 AA.
 AC Q9H6P7; 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE CDNA: FLJ22005 FIS, CLONE HEP06902.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
 RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
 RT "NEO human cDNA sequencing project."
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK025658; BAB15207.1; -
 SO SEQUENCE 375 AA; 42401 MW; 97A51B0140F9DD09 CRC64;
 Query Match 31.3%; Score 1953; DB 4; Length 375;
 Best local similarity 99.7%; Pred. No. 5, 6e-94;
 Matches 374; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 840 MSKLEGNLEPPDLEETLLOEPSTGNKDESLPTITDVISSEKKEKQTSADRTTIOE 899
 |||||||
 DB 1 MSKLEGNLEPPDLEETLLOEPSTGNKDESLPTITDVISSEKKEKQTSADRTTIOE 60
 |||||||
 QY 900 NEYSEGVSYVDQSLDIHIEGCTNDSQSKCDYKSVOPPEPHKVVHSEHLNLPVOY 959
 |||||||
 DB 61 NEYSEGVSYVDQSLDIHIEGCTNDSQSKCDYKSVOPPEPHKVVHSEHLNLPVOY 120
 |||||||
 QY 960 SVOCSPESFAFRSHSLPPKKNKNSLLIGSLGLDANNPKMLRTCSLPDLSKLEFRL 1019
 |||||||
 DB 121 SVOCSPESFAFRSHSLPPKKNKNSLLIGSLGLDANNPKMLRTCSLPDLSKLEFRL 180
 |||||||
 QY 1020 MDVPTGVADYRQDNLDEIDELIDENIKESPSDSBDIVFEETDTDLOELQASMEQLLRQEPGE 1079
 |||||||
 DB 181 MDVPTGVADYRQDNLDEIDELIDENIKESPSDSBDIVFEETDTDLOELQASMEQLLRQEPGE 240
 |||||||
 QY 1080 EYSESEESVYLNKNSDVEPTANGTDVADEDNPPSSSALNEMHSDNSNGEJASSECDSVF 1139
 |||||||
 DB 241 EYSESEESVYLNKNSDVEPTANGTDVADEDNPPSSSALNEMHSDNSNGEJASSECDSVF 300
 |||||||
 QY 1140 NHLEELRLHLEQEMGEKFEFFVEYKIKAIHEDDENIEICISKYONITLGNHGHOLYAKITL 1199
 |||||||
 DB 301 NHLEELRLHLEQEMGEKFEFFVEYKIKAIHEDDENIEICISKYONITLGNHGHOLYAKITL 360
 |||||||
 QY 1200 HLMVADGAYOEONDE 1214
 |||||||
 DB 361 HLMVADGAYOEONDE 375
 |||||||
 RESULT 4
 Q96S52 PRELIMINARY: PRT: 320 AA.
 ID Q96S52

AC Q96S52; 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE CDNA: FLJ14674 FIS, CLONE NT3P2003912, HIGHLY SIMILAR TO
 DE SERINE/THREONINE-PROTEIN KINASE NEK1 (EC 2.7.1.-).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
 RA Matsumura H., Ishii S., Kawai Y., Salto K., Yamamoto Y., Wakamatsu A.,
 RA Nakamura Y., Nagahara K., Masuh Y., Sasaki N.;
 RT "NEO human cDNA sequencing project."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK027580; BAB55209.1; -
 SO SEQUENCE 320 AA; 36509 MW; F64FE9BB69C0F730 CRC64;
 Query Match 22.2%; Score 1387; DB 4; Length 320;
 Best local similarity 100.0%; Pred. No. 1, 1e-64;
 Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 208 MNVLVYKTIISGSEFPVSYHYSDIRSLVSQLFKNRPDRPSVNSILEKGFIAKRIEKL 267
 |||||||
 DB 1 MNVLVYKTIISGSEFPVSYHYSDIRSLVSQLFKNRPDRPSVNSILEKGFIAKRIEKL 60
 |||||||
 QY 268 POLIAEEFLTKFKFSGQPIPAKRPASGONSISVMPAQKITRPAPKYGIPLAYKKYGD 327
 |||||||
 DB 61 POLIAEEFLTKFKFSGQPIPAKRPASGONSISVMPAQKITRPAPKYGIPLAYKKYGD 120
 |||||||
 QY 328 KLEHKKPLOKHQAOHPTEKRVNNGEERKISEAAKRRLEFEKKEKKOKDQIISLMA 387
 |||||||
 DB 121 KLEHKKPLOKHQAOHPTEKRVNNGEERKISEAAKRRLEFEKKEKKOKDQIISLMA 180
 |||||||
 QY 388 EOKKROEKRLERINRARGGRNVLASGSEYKAPFISGGTIAVSSSSSGOYEYH 447
 |||||||
 DB 181 EOKKROEKRLERINRARGGRNVLASGSEYKAPFISGGTIAVSSSSSGOYEYH 240
 |||||||
 QY 448 AIFDQMOQOQRAEDNEAKKRELYGGLPER 477
 |||||||
 DB 241 AIFDQMOQOQRAEDNEAKKRELYGGLPER 270
 |||||||
 RESULT 5
 ID Q95673 PRELIMINARY: PRT: 744 AA.
 AC Q95673; 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE MSTRK25 KINASE-LIKE PROTEIN.
 GN NEK4 OR MSTRK25.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-BALB/C;
 RA MEDLINE:20001940; PubMed-10529384;
 RA Hayashi K., Igayashi H., Ogawa M., Sakaguchi N.;
 RT "Active and substrate specificity of the murine STRK2
 RT regulator protein NIMA of Aspergillus nidulans."
 RL Biochem. Biophys. Res. Commun. 264:449-456(1999).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: Y09234; CAI0436.1; -
 DR HSSP: 063450; 1A06.
 DR MGI:134404; Nek4.

Wed May 15 14:14:06 2002

us-09-783-320-4.rspt

Page 4

DR InterPro: IPR0000719; Euk_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR Pfam: PF00069; pkinase_1.
DR SMART: SM00220; S_TKc_1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_STY_1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase
KW Sequence 744 Aa; 83542 MW; 82ELAAFE99CC72A CRC64;

Query Match	Similarity	13.4%	Score 836	DB 11	Length 744
Best Local	Similarity	30.1%	Pred. No. 1,3e-35		
Matches	226	Conservative	128	Mismatches	261
				Indels	160
				Gaps	24
QY	4	YVNLQKIGSGSREKALIVSTEDGQYQYIKKINISRRSSKREKRESREKAVLANMKHPN	63		
Db	6	YCYHRRVAGKSGTGEVTLVHNRDGGQYQYIKKINLNKNSRRERRAADEAOLLSQLNHPN	65		
QY	64	VOYRESEF-BNGSLYIVMDYCEGDLFEKFINAQGVLFQEDQIIDMFVQICALKVHNR	122		
Db	66	VYTKESWEGDGLIYIVMGFCGGDLRYKKEQKQGLLPESQVVEWVQITAMLOYLTHK	125		
QY	123	KILHARDIKSONIFLTKDGTVOJGDGIGIARVANSYVELARTICIGPYPLSELIEENRPN	182		
Db	126	HIHLRDLKQONVLEFTNIIKIVGDGIGIARVLENNHGDMASTLIGHPYMSBELTSNKNPN	185		
QY	183	KSDIMALGCVLELCLKHAFAGSMKMLVAKIISGSPRPVSLYSVDLSLSVQSOLFKN	242		
Db	186	KSDVMAAGCCVYEMALTKHAFNAKMNSLYRIIEGKLPMPKPYSTLELLIRTMISLR	245		
QY	243	PRDRPSVNSLIEKGFATRIEKFSLPOLIAEFCLKTFPSKSG-----SOPIPAKRPAS	295		
Db	246	PEBRPSVSLIRQPYIKHISLFL-----EATRAKT-SKNNVKNGDSRAKPAVAAYVSR	298		
QY	296	GQNSISVMAQKTTKRAAKYGIPLAKYKKGKQKILHEKPPLO-----KKHQAHPREK	348		
Db	299	EESNTDVIHQ-----PRSEGSAL-HWAGEDKLSQEKRPVIGIRPASPALBEGHTKQDM	353		
QY	349	VYNGEERRKRISENAKRRLEFIEKEK-----QKQDILSLMKAEQMK-----ROEK	395		
Db	354	NMGESCATIS-----RINIDILPERKDSANAGVQDSQPHADADVEDVQSCSISOEK	408		
QY	396	ERLERINRAEQGWNRVL-----SAGSG-----EVKAPFLGSGGTIAPSSFSKQO-YEHY	446		
Db	409	ERLDGNTKTSDDQ-GNLPRRSSDGDGDEGSELVKKPLYPNSKDQKPPQFOELRPLRLPSY	467		
QY	447	HAI-FDOMQOQRAEDNE-----AKMKREIYGRGLPRQKQOLAVR	486		
Db	468	PGJGKVDIATQONDNDGPGVAGCVNSSRTSTASAK-----DRPLSARER-----RR	516		
QY	487	AKQVEEFLQKREKRAMONKRAAGHHVYLAIRQLQIFNE-----ROOTIKK	534		
Db	517	LKQSOEEMLPSPGPAVQRTPSAVEPLKPOEDQPIPAQFSSDCSITQMNHTLPREREKRL	576		
QY	535	LRG-EKKKANHSQCGGSEADMRKKKILSKAHANNAAMALTKQOLRKKKRAYE-----	588		
Db	577	MHLSDDELSSSTSSITDKSDGDSRG-----KSHINNEKDLVQMLTQTLLEAKESCEDL	631		
QY	589	-----REKVMYEEHLVYAKGVKSSVDYSPPLQCHETGSGSPSKQDMRSVISTALKEY	639		
Db	632	QVNLPGSEFRLHAKITDYLILKGAKEEVEPRCTELPGLIIPSEKILRRIRIVEVRA-----	687		
QY	640	GVDSSLDTRETISEMQKTNNAISSKREILRLRLLENLKAODEKGMONLSDFEIVNHED	699		
Db	688	-----DVYIGL-----GIOLLEQVDFDLGEBD	709		
QY	700	AKHEH 704			
Db	710	ELERE 714			

RESULT 6

Q9R1J1	PRELIMINARY;	PRT;	792 AA.
ID Q9R1J1			

```

01-MAY-2000 (TREMBlrel. 13, Created)
01-MAY-2000 (TREMBlrel. 13, last sequence update)
01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE SERINE/THREONINE-PROTEIN KINASE NEK4.
GN NEK4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID:10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:99321807; Pubmed:10393247;
RA Chen A., Yanai A., Arama E., Kilfin G., Motro B.;
RT "NIMA-related kinases: isolation and characterization of murine nek3
RL Gene 234;127-13(1999).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC EMBL: AF099067; AAD16287.1; -.
DR HSSP: Q63450; 1A06.
DR MGD: MGI:1344404; Nek4.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR SMART: SMO0220; S_TKc_1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE: PS3011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS0108; PROTEIN_KINASE_SF; 1.
SW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 792 AA; 89093 MW; 9B8F37311377CC6 CRC64.

```

Query Match	Similarity	13.4%	Score 836	DB 11	Length 792
Best Local	Similarity	13.4% <td>Pred. No. 1.5e-35</td> <td></td> <td></td>	Pred. No. 1.5e-35		
Matches 239	Conservative	129	Mismatches 269	Indels 186	Gaps 22
Oy	4	YVRLKIGESPEKALIVSTEDGRQYVYKINIRSMRSKREERSREVVANMMKHP	63		
Db	6	YCYKRVVGRSSTGEYVLVKKHRDGGKOYVYKIKNLNMASSRRERRAAEDQALLSQLKHPN	65		
Oy	64	VQYNESE-ENGSLTYMDYCBGGDLFKRNQKGVLEQDQILDMVQJLAKVHYH	122		
Db	66	VTVSEGGGGLTYVGMGCEBGGDLTKYKQKQKQLPSEQVEMVYQALMQLYHEK	125		
Oy	123	KILHRDIKSONIFLTKDGYVQJGDFGIARVNLSTYELARTCIGPYTLSPICEKNPYNN	182		
Db	126	HLHRDKTONVLETKRNIIKVGDDIAVLENNHDMQASTLIGPYTMSPELSNRKYN	185		
Oy	183	KSDIWAQCVLYELCTLKAHFAEAGSKRNLYLKIISSFPVSLHYSTDLNLSVQLFKRN	242		
Db	186	KSDYWAQGVCCVEYEMATLKAHFNKDNMNSLYVYIEGKLPMPKPVYSTELMELRTMLSRK	245		
Oy	243	PRDRPSVNIILEKGFIAKRIEKLSPQLIAEEFLCKTFSKFG	295		
Db	246	PEERPSVRSITLRPHYIKHHSIFL	298		
Oy	296	GQNSISVMPQKITTTPAKAYGIPLAVKAYGDKRIHEKKRPD	348		
Db	299	EESNTVDIHHQ	353		
Oy	349	VNTQEEERKISEBAARRRLEFEIEKK	395		
Db	354	NNTSESCATIS	408		
Oy	396	ERLEIRINAREEQWRVNL	448		
Db	409	ERLQGNKTSDDOP	467		
Oy	449	IFPDQMO	476		

RT	mouse and in zebrafish.	EMBL/GenBank/DBJ databases
PL	Submitted (Aug-2001) to the	
DR	EMBL: AF407579; AAL09675.1;	
KW	kinase	
SQ	SEQUENCE 698 AA; 75264 MW;	686B29A8CF180E94 CRC64;

[illegible]

RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shindo S., Takuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.:
"Sequence and analysis of chromosome 3 of the plant Arabidopsis
thaliana".
RL Nature 408:820-823(2000).
DR EMBL: AC009665; F54AG542.1. "-."
DR HSSP: Q6450; A00719: Euk.pkinase.
DR InterPro: IPR002685; ppr.
DR InterPro: IPR002280; Ser_thr.pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR InterPro: P00069; pkinase.1.
DR Pfam: SM00220; S_TKc.1.
DR SMART: SM00219; TYRKc.1.
DR PROSITE: PS00107; PROTEIN KINASE_ATP; UNKNOWN-1.
DR PROSITE: PS00011; PROTEIN KINASE_DOM.1.
DR ATP-binding: kinase; Transferase.
KW SEQUENCE 606 AA; 68034 MW; 1873372C6CE3B9C CRC64;

Query Match	10.0%	Score 621.5	DB 10	Length 606
Best Local Similarity	27.6%	Pred No. 14e-24		
Matches 177	Conservative 121	Mismatches 237	Indels 107	Gaps 19
1	MEKYRVLQRIEGSGFAGLALVYKSTEDGRQYVIREINISMSKREERERRAVIAMKH	60		
1	MEYVELLEIDGSGSGLVLRHNEHKKLVYKLIHQAQTRTRSAHQEMELISKIHN	60		
1	PNIVQRESFESENGS-LYIVMDYCEGDDLFRKINAOQVLPDEDDILDMFOICLAKHV	119		
61	PFVEYKSDSWVKKGYCVCIITIGYCKGDMALAKTNGVHFEKIKCMVQIILALLEYL	120		
120	HDRIILAHDIKSONIFLTKDSTVOLGDFGFLARVLNSTVELARFCIGTPYUSPEICENPK	179		
121	HAHHIILHRYVCKSNIFLTKDDIRLGDGLAKVLTSD-DLASSVGTPEYCMPELLADIP	179		
180	YNNKSDIWAIGCVLYELCTLKHAFEGASMKNLVLTITSGSPFVSLHYSYDLRLSIVOLF	239		
180	YGSQSDIWSIGCCMYEMTAMKPAFAFMQGLINIRKSTVPLPRLQYGAARGLYKSWL	239		
240	KRNRPDRPSVNSITLKGFLAKRIEKFELSPDILAEFCLETFKSFQSPVPAKRPASGQMS	299		
240	RKNLELPSAELLRPLDLPYIOL-----THLKV-NDGSIVLPAQMPSEESAR	288		
300	ISVWPAKIKTRPAKAGIPLAVYKKYGDKKLHEKKPLQKNHOAQOTPEKRVNTEERKITS	359		
289	RNSPEQDR-RRPAAK-----SHSFGSPSRGMLF-----	316		
360	EEAAKRRLEPIEKKOKDOIISLKAQOMKROEKERLEIRINRAEQGMRYVLSAGSG	419		
317	-----DSVSISIKTVP-----AYLNRROVD-----LSTDASG	344		
420	EYKAPFLGSGTIA--PSSFSRQOEYHVAIFDQV--OQOFAEDNEAKMKRELYGR---	472		
345	D-----GVVRRRSEAKSRYSRYVPVAVASAPVRPQPSDSDGOLVSSOLKMKRA	395		
473	GLPEROKQGLAVRKQOEEFLQKRPAMONKARAGAHNVYLAIRLOIR--LQNFENROQ	530		
396	ALTRPASMSSSKRAKELKDSLYLSTKSFJHOINSIDVSMNAPRIDKLEPFLASTEERF	455		
531	IKAKLNGEKKK-----NHSQEGSEEDAKRKIKTSLKANHANAARAAVLEKOLEKKREA	586		
456	VPV-VGKKKKKSSSGSPPREPLDOSTIKDF--TLEGQGNRBDALIMKAVYE---EDA	510		
587	YERKKVWEHLVAKGVYSDVSPRLQGHETGGSPSKQOYRS	628		
511	YLEDRSESSDNATAGASSRASS--GVRRQRFPPSSYOORA	549		

AC Q90XC2: 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE NIMA-RELATED KINASE 8.
 GN NEK8.
 OS Brachydanio rerio (zebrafish) (Zebra danio).
 OC Actinopterygii: Neopterygii: Teleostei: Euteleostei: Ostariophysi:
 CC Cypriniformes: Cyprinidae: Danio.
 OK NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Liu S., Lu W., Obara-Ishihara T., Drummond I., Beier D.R.:
 RT "A defect in a novel Nek-family kinase causes cystic disease in the
 mouse and in zebrafish."
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF407580; AAL09676.1;
 KW kinase.
 SQ SEQUENCE 697 AA; 76541 MW; 875A51D1F3831AA7 CRC64;

Query Match 9.9%; Score 617; DB 13; Length 697;
 Best Local Similarity 34.5%; Pred. No. 2,9e-24;
 Matches 169; Conservative 59; Mismatches 140; Indels 102; Gaps 11;

OY 1 MEKVRLOKIGEGSGFKAIIYKSTEDGRQYVKEINISMSSKEEBSREAVANLAMKH 60
 DB 1 MEKTERLRVGRGAFGLVHLCRRRTSAVILKEIPEDMTREDAQNECOYLKLLSH 60
 OY 61 PNIVQYRESFEENGSLIYVNDYCEGDLFRINAQGVLFQEDQILDMFOVQICLALHNV 120
 DB 61 PNIEYYENFLEDKALMAIEMVAPGGLADYDQKRNNSLDLDDTLHSFVQILLALHNV 120
 OY 121 DRKILHNDIKSONIFLTR-DGVQLDQFIAVLNLSYELARTCIGTPYVLSPEICENK 179
 DB 121 NKLLHNDLKTQNLILDKHQMIVKIGDFGISKILVSKS-AVTYVGTFCYISPEICGPK 179
 OY 180 YNNSSDIWAIGCVYELCTLKHAFEGSMKNVLTIGSPPPVLSHYSLRSLVSOLE 239
 DB 180 YNKSDDIWAIGCVYELASLKRAFEAANPALVYLKIMSGTFAPISDRSPELRLIIML 239
 OY 240 KRNPRDRSVNSILEKFI-----AKRIEKLSPDLIAEFECCLATFEKFG 284
 DB 240 NLDSKRQQLNEIMAHICIRPLINLYTDIGVKKRIEKLPS-----N 283
 OY 285 SQPIPAKRPASGNSISVPAOKITPAKYGIPLAAYKKYGDKKLHEKKPLQKHQAHOT 344
 DB 284 VQAGPHGHPGWM-----ITSTRTRGGL-----SSLTSSKMMHPL 317
 OY 345 PEKRVNTEGERRKISEBAARRLEFIEKKQKQDOILSLMAEQMKROEKRLERINRA 404
 DB 348 PLFYSYTW-----GSGISTPLRLPMINTEVIO-----VSLGRQKMGYTKSGRLIT 363
 OY 405 REQGMRVNLSAGSGEYKAPFLSGGTLAPSSFSRSGCYEHNAIFDDMOQO 456
 DB 364 -----W-----EAPSVSGEPTLPGA-----VEQMQO 386

RESULT 13
 O9D685 PRELIMINARY; PRT: 291 AA.
 AC Q9D685: 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
 DE 4632401F23RIK PROTEIN.
 GN 4632401F23RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
 CC Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae; Mus.
 OK NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6J; TISSUE=SKIN;
 RX MEDLINE=21085660; Pubmed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Araiwa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kusukawa T., Saito R.,
 RA Kadoya K., Matsuda H.A., Ashburner M., Batilov S., Casavante T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
 RA Kuehl P., Lewis S., Matsuo T., Nitaoid I., Pesole G., Quackenbush J.,
 RA Schiraldi L.M., Scuderi F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Bonfieldi D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RT Nature 409:685-690(2001).
 CC -I- SIMILARITY: BEYONDS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AK014545; BAB29424.1; -
 DR MCD: MGI:1921569; 4632401F23RIK.
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR002290; Ser_Thr_Pkinase.
 DR InterPro: IPR001245; Tyr_Pkinase.
 DR Pfam: PF00069; Pkinase.
 DR PRINTS: PF00109; TYRKINASE.
 DR SMART: SM00220; S_TKc; 1.
 DR SMART: SM00219; TYKc; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 291 AA; 32802 MW; DF2B752B4467021F CRC64;

Query Match 9.8%; Score 614; DB 11; Length 291;
 Best Local Similarity 42.4%; Pred. No. 1.4e-24;
 Matches 129; Conservative 56; Mismatches 93; Indels 26; Gaps 5;

OY 1 MEKVRLOKIGEGSGFKAIIYKSTEDGRQYVKEINISMSSKEEBSREAVANLAMKH 60
 DB 1 MEKTERLRVGRGAFGLVHLCRRKADQKLVILKQIPEDMTREDAQNECOYLKLLSH 60
 OY 61 PNIVQYRESFEENGSLIYVNDYCEGDLFRINAQGVLFQEDQILDMFOVQICLALHNV 120
 DB 61 PNIEYYENFLEDKALMAIEMVAPGGLADYDQKRNNSLDLDDTLHSFVQILLALHNV 120
 OY 121 DRKILHNDIKSONIFLTR-DGVQLDQFIAVLNLSYELARTCIGTPYVLSPEICENK 179
 DB 121 THLLHNDLKTQNLILDKHQMIVKIGDFGISKILVSKS-AVTYVGTFCYISPEICGPK 179
 OY 180 YNNSSDIWAIGCVYELCTLKHAFEGSMKNVLTIGSPPPVLSHYSLRSLVSOLE 239
 DB 180 YNKSDDIWAIGCVYELASLKRAFEAANPALVYLKIMSGTFAPISDRSPELRLIIML 239
 OY 240 KRNPRDRSVNSILEKFIKRIEKLSPDLIAEFECCLATFEKFG 284
 DB 240 SLEPAGPPLSHW-----AOPLCIALALNHITDVG--VMMKRPVQ 279
 OY 296 GONS 299
 DB 280 GDGS 283

RESULT 14
 O94C05 PRELIMINARY; PRT: 943 AA.
 AC O94C05: 01-DEC-2001 (Tremblrel. 19, Created)

Wed May 15 14:14:06 2002

us-09-783-320-4.rspc

Page 10

```
OY 297 QNSISVMPAKITTPAKKYGILAYKXYGDKKHLHEKKPLOKHOAHQTPKRVNTGEER- 355
Db 283 DNDV-----KKTIFA-----ISESTPVKKNR-----EKRMSCGNDRT 314
OY 356 -----RKISEA--ARRRLEFIEKE 374
Db 315 LNPVSDDHDTFENRRYPKTPSRVSELVSGSPDRGSTVTKKITSKALLVNNPQIVPKL 374
OY 375 KKKODIISLMKAEQKKROEKER-----LERINRAREQWNRNVLASGSGEVKAPF 425
Db 375 TTPPARQVLDLRNSDMASFTLVKRSVSTTRASLPLTNKAAYQELPRRPSLSFLDCIKSPD 434
OY 426 LSGSGTIAPS-----SESSRGQYEHYHAFDQMGOORADNEAKKREIYGRGLPERQKG 480
Db 435 VSYN---APRIDKMLEFFPLASTYEDPFHP-----RTSSNSA-----QSSSGSPOA 477
OY 481 QLAVERAKOVEEFLORRREAMONK---ARAGSHVYTLARLOIRLONFNERQOIKAKLRG 537
Db 478 EYSVMKDKCTIOLPSKFDPMSSNDAMOGYEGPMVHV-----DREDIT----- 520
OY 538 EKKRANHSEGOEGSEADMRKK--TESLKAHANARAALKEQLEKRRKREAYE----- 588
Db 521 DSSDONATAGASSRTSSDTRRRRRFNNSSYKQRAEALLEGLEFSARLLQEDRFDELGVLLK 580
OY 589 -----REKKVWEHVLVAKGVKS 606
Db 581 PFGPEKVSPPRETAIW-----LTKSIKEN 603
```

Search completed: May 15, 2002, 08:01:42
Job time: 389 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 15, 2002, 07:51:30 ; Search time 56.63 Seconds
(without alignments)
2059.906 Million cell updates/sec

Title: US-09-783-320-4
Perfect score: 6243
Sequence: 1 MEKYVRLQKIGESFGKAIL.....YAKILHLVMDAGVQEDNDE 1214

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR-71:
1: PIR1:
2: PIR2:
3: PIR3:
4: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3385	54.2	774	2	S25284	protein kinase nek
2	847.5	13.6	841	1	T78885	serine/threonine-s
3	836	13.4	792	2	UC7122	protein kinase (EC
4	565.5	9.1	431	2	T11854	protein kinase-lik
5	557	8.9	941	2	T49136	NIMA-like protein
6	550	8.8	445	2	G01452	hypothetical prote
7	540.5	8.7	357	2	T29771	probable protein k
8	534.5	8.5	699	2	A43734	probable G2-specif
9	514.5	8.2	722	2	T37970	GIN4 protein - yea
10	502.5	8.0	1243	2	S59359	serine/threonine p
11	500	8.0	1233	2	T14157	NIMA-like protein
12	499.5	8.0	779	2	A57177	hypothetical prote
13	475	7.6	200	2	B96587	protein kinase - s
14	467	7.5	1094	2	S49313	protein kinase SK2
15	464.5	7.4	1206	2	T34021	hypothetical prote
16	462	7.4	1211	2	T27522	hypothetical prote
17	455	7.3	1231	2	T18532	serine/threonine s
18	455	7.3	1354	2	S74244	serine/threonine s
19	452	7.2	1518	2	S37928	probable purine nu
20	451	7.2	1233	2	T30989	serine/threonine p
21	451	7.2	1558	2	T29253	hypothetical prote
22	445	7.2	1354	2	S69211	serine/threonine s
23	437.5	7.0	915	2	S74381	probable protein k
24	432	6.9	435	2	S23580	serine/threonine s
25	427	6.8	1388	2	S74245	serine/threonine s
26	425	6.8	1388	2	S70633	serine/threonine s
27	422	6.8	1732	2	T14039	protein kinase (EC
28	421.5	6.8	1895	2	T15881	hypothetical prote
29	418.5	6.7	294	2	T21075	hypothetical prote

30	407.5	6.5	925	2	A55748	protein kinase (EC
31	405	6.5	1001	2	T17365	serine/threonine p
32	402	6.4	740	2	T24340	hypothetical prote
33	402	6.4	797	2	T23927	hypothetical prote
34	400	6.4	1702	2	T14050	protein kinase (EC
35	399	6.4	1114	2	T18224	protein-serine/thr
36	398.5	6.4	1245	2	G89287	hypothetical prote
37	393	6.3	1080	2	S48944	hypothetical prote
38	393	6.3	1173	2	T25539	probable serine/th
39	390.5	6.3	1192	2	T18611	probable serine/th
40	388	6.2	1062	2	S46367	hypothetical prote
41	387.5	6.2	1228	2	T18987	hypothetical prote
42	383.5	6.1	465	2	B57748	protein kinase (EC
43	383.5	6.1	705	2	A48144	protein kinase CDC
44	382.5	6.1	740	2	T18536	ribosomal protein
45	382	6.1	713	2	S27966	probable serine/th

ALIGNMENTS

RESULT 1
S25284
protein kinase nek1 (EC 2.7.1.-) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Sep-1999
C:Accession: S25284
R:Metzger, K.; Mizzen, L.; Motro, B.; Ben-David, Y.; Bernstein, A.; Pawson, T.
EMBL J. 11, 3521-3531, 1992
A:Title: A mammalian dual specificity protein kinase, Nek1, is related to the NIMA ce
A:Reference number: S25284; M01D:93010942
A:Accession: S25284
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-774 <DB>
A:Cross-references: GB:S45828; NID:g256854; PIDN:AAB23529.1; PID:g256855
C:Gene: nek1
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
C:Keywords: ATP; phosphotransferase
F:2-258/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif

Query Match	Similarity	Score	DB 2:	Length	774:
Best Local	85.5%	Pred. No. 5e-96	64:	Indels	4:
Matches	665:	Conservative	45:	Mismatches	64:
				Gaps	3:
OY	1	MEKYVRLQKIGESFGKAILVSTEDGROYVYIKETINISRMASRRERREAVIANMKH	60		
DB	1	MEKYVRLQKIGESFGKAVLVSTEDGRHYVYIKETINISRMASRRERREAVIANMKH	60		
OY	61	PNIVQRESEFENGSLYIVMDYCEGGDLFRINAQGVLFQEDDILDMFYQICALKVH	120		
DB	61	PNIVQRESEFENGSLYIVMDYCEGGDLFRINAQGVLFQEDDILDMFYQICALKVH	120		
OY	121	DRKILHRIKSONFFLRKDGTVOLGDFGIARVNSTVETLARTIGTPYYSPEICENKY	180		
DB	121	DRKILHRIKSONFFLRKDGTVOLGDFGIARVNSTVETLARTIGTPYYSPEICENKY	180		
OY	181	NNKSGIVMLGCVLVECTLKHAFFASMKNLVLTISGFFPVSLHYSDRLSLVOLF	240		
DB	181	NNKSGIVMLGCVLVECTLKHAFFASMKNLVLTISGFFPVSPHYSDRLSLVOLF	240		
OY	241	RNPDRPVSNTLTKGFLAKRIEFLSPOLIAEEFLTKFSGSQPIPARPAGGNSI	300		
DB	241	RNPDRPVSNTLTKGFLAKRIEFLSPOLIAEEFLTKFSGSQPIPARPAGGNSI	300		
OY	301	SWPMAQITTKPAKYGPILAYKKYGDKKLHKKFLDKHQAQOTPEKRVNNGEERKISE	360		
DB	301	SWPMAQITTKPAKYGPILAYKKYGDKKLHKKFLDKHQAQOTPEKRVNNGEERKISE	360		
OY	361	EAARRRLEETLEKKKQDIIISLMKAEQKKROEKRELEINARQGGWRNVLSAGSGE	420		

```

Db 361 EAAKRRRLERLEKRRKKOD--IRFLAEOKKROEKORLERLNKADQGRNRLRGGSGE 419
Oy 421 VKAPFLSGGTIAPSSFSRSGOYEHYHAIFDMOQOORADENAKMKREITGRLPEROG 480
Db 420 VKASFFIGGAVSPSPGQYEHYHAIFDMOQLRADNENARKGGYGHMLPEROGK 479
Oy 481 QLAVERAKOYEELQORREAMONKARABAGHMYVLARLROIRLONFNEROQIKAKLGEKK 540
Db 480 HLAVERANOVVEELQORREAMONKARABAGHMYVLARLROIRLONFNEROQIKAKLGEKK 539
Oy 541 EANSSEOGESSEADMRKKIESIKAHANNAARAVLKEQERKKREKVEERKVMVEHLYA 600
Db 540 EADGTQGEATEETEDMRKKMESLKAQTNAARAVLKEQERKKREKVEERKVMVEHLYA 599
Oy 601 KGVSPGVSPTLQGHETGSGSPKQOMRVSYSALKVEGVDSLTDPRETSEEMOKTNN 660
Db 600 R-VKSSDVPRLPELLETGSGSPKQOYKPVYSTALKELGDSLTDOE--EENKESNS 656
Oy 661 AISKREILRLNENLKAODEKGMQNLSDTEINVEDAKHEKEKSVSDRKKWEKGG 720
Db 657 AISKREILRLNENLKAODEKGMQNLSDTEINVEDAKHEKEKSVSDRKKWEKGG 716
Oy 721 QLVIPDELDTLDTSTFTEHRTVGEVYIKLGPNGSPRRANGKSPDTSVLKILGEAELOL 778
Db 717 QLVIPDLAVLDTSTFTEHRTVGEVYIKLDSNGSPRRANGKSPDTSVLKILGEAELOL 774

```

```

RESULT 2
17885
serine/threonine-specific protein kinase (EC 2.7.1.1) SK2 - human
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 29-Sep-1999
C:Accession: 178885
R:Levedekou, E.N.; He, M.; Baplist, E.W.; Craven, R.J.; Cance, W.G.; Welch, P.L.; Simic
O:Oncogene 9, 1977-1988, 1994
A:Title: Two novel human serine/threonine kinases with homologies to the cell cycle regu
A:Reference number: 158396; MID:9426888
A:Accession: 178885
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-841 <RES>
A:Cross-references: GB:L20321; NID:9348244; PIDN:AAA36658.1; PID:9348245
C:Genetics:
A:Gene: GDB:SK2
A:Cross-references: GDB:374125
A:Map position: 3p21.1-3p21.1
C:Superfamily: human serine/threonine-specific protein kinase SK2; protein kinase homo
C:Keywords: phosphotransferase.
F:4-261/Domain: protein kinase homology <KIN>

```

```

Query Match 13.6%; Score 847.5; DB 1; Length 841;
Best Local Similarity 28.1%; Pred. No. 3, 1e-19;
Matches 263; Conservative 160; Mismatches 316; Indels 197; Gaps 34;
Oy 4 YVRLOKIGESFGKALIVKSTEDGRQYVKEINISMSSKEKREKREYAVLANKKHNT 63
Db 6 YCYIRLVGKSGYEVTLTKRRDQKQVYIKLKLNNASSRERRAAEQEQLSOKHNT 65
Oy 64 VOYRESEF-ENGSLYIVMDYEGGDLFKRINAKGVLFQEDQILDMFVOICLAKHVR 122
Db 66 VTYRESEWGGGLLYIVMGCEGGDLYRLKLEQKQLLEPNQVVEFVQIAMAQLYHEK 125
Oy 123 KILHRDKSONIFLTKDGTQLDGFARVLNSTVELATCTGTPYLSPELCEKPNNN 182
Db 126 HILHRDKTONVFLTRNLIKVDLGIARVLEHNCMASTLGTPTYMSPELFSKPNY 185
Oy 183 KSDIHALGCVYELCTLKAHFEAGSMKLVKTIISGSPVSLHYSYDLSRISYDLFRN 242
Db 186 KSDVWALGCCVYEMATLKAHFNKDKMSLVYRIIEGKLAMRDRDSEPELAEILIRMLSK 245
Oy 243 PRDRPVSNIIEKFIKRIEKLFLQLAEFCCLTFSTKFSQSLPAPKRRPASGONSISV 302

```

```

Db 246 PERRSVRSILKOPYIKQIISFLE-----ATIKSKNNT-- 281
Oy 303 MPKQKITKPAANYGIPLAVKRYGDKLHEKKPL-----QKHQAQOTPEKRVNTE 353
Db 282 -----KNGDS--QSKPFATVVGAEASNHEVIRHP----- 310
Oy 354 ERKISEEAKRRRL--EFIEKKKKQDQI--SLKAEQMKRQ-----EKERLERIN 402
Db 311 ----LSSEGSQYIIGEEKCLSOEKPRASGLKSPASLKATKODISNTTELATISYN 366
Oy 403 RAREQGNRVLSAGSGEYKAPFLGSGGTIAPSSFSRSGOYEHYHAIFDMOQOORADE 462
Db 367 -----IDLPAKRDVSDGFQEN--QPRLDASNLGIGCSI--SQVEEEMQDNT 415
Oy 463 AK-----WKREYGRGLPEROGK--QLAVRAKOVEEFLQORREAMONKARABAGM 511
Db 416 KSSAQEPNLIIPWSSDIYTGKNEPVKPLQLEQKPKQSL-----ALSPKLESGTI 470
Oy 512 VYLARLROIRLONFNEROQIKAKLGEKKEANHSEGGSEEDMRKKTE--SLKAHANA 570
Db 471 LAHSNL--RLGSSDPSASRAVAGITGVCHHADDVAGECTIEKGRHPDLQPHNSG 527
Oy 571 RAAVLKEQLERRKREYERKKVMEHLVAKVKSQVSP--LQGHETGSGSPSKQOMS 628
Db 528 SEPSLSRORRORRROTHER--GEKKQVRRLDLPAPQESPPRLPSPHIVGK----- 576
Oy 629 VISVTSALKEVG----VDSLTDTRETSEEMOKTNNAISSKREILRLNENLKAODEK 683
Db 577 -VDVTSQKEANORRVYTGVSISR--SEMSSSKDRPLSARE--RRRLKO--SOEEMS 629
Oy 684 GMQNLSDPFEINVEDAKEHEKEK-----SVSSDRKKWEAGQLVLPDLDTLD 732
Db 630 SSGPSYRKASLSVAPBGRPOEDDQPLPARLSQCVSTQKQHC-----LSEDLSSS 684
Oy 733 TSPSTTERHTVGEVYIKLGPNGSPRRANGKSPDTSVLKILGEAELOLOTELL-----E 784
Db 685 TSTSDKSDQDYE-----QKGTQNEI--NALVQMTQTLKIDSESE 725
Oy 785 NTINSEIS--PEGKYRPLITGEKKVOCISHEIN--PSAIVDSFVETKSPSESA-- 836
Db 726 DVPANPVSFEFLKRRKYDITLHGKVAEEAEELHEKELPSALM--PGSEKIRRLVEYLR 783
Oy 837 -----SPQMSLKLEGNLEPDDLEEI--LQEPSC 864
Db 784 TVIRIGLGVQLLEQVYDLLEEDFEFRRLREHMG 819

```

```

RESULT 3
JC7122
protein kinase (EC 2.7.1.37) 2 - mouse (strain balb/c)
N:Alternate names: serine (threonine) protein kinase
C:Species: Mus musculus (house mouse)
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 31-Mar-2001
C:Accession: JC7122
R:Hayashi, K.; Igarashi, H.; Ogawa, M.; Sakaguchi, N.
O:Biochem. Biophys. Res. Commun. 264, 449-456, 1999
A:Title: Activity and substrate specificity of the murine SK2 serine/threonine kinas
A:Reference number: JC7122; MID:20001940; PMID:10529384
A:Accession: JC7122
A:Molecule type: mRNA
A:Status: Preliminary
A:Residues: 1-792 <HAY>
A:Cross-references: GB:AJ223071; NID:94138208; PID:94138209
C:Genetics:
A:Gene: MSTR2L
C:Superfamily: human serine/threonine-specific protein kinase SK2; protein kinase ho
C:Keywords: phosphotransferase

```

```

Query Match 13.4%; Score 836; DB 2; Length 792;
Best Local Similarity 29.0%; Pred. No. 6, 6e-19;
Matches 238; Conservative 129; Mismatches 269; Indels 186; Gaps 22;

```

[illegible][illegible]

```

OY 240 KRNPRDRPSVNS-----ILEKGI AKRIEKLISPOLIAEFLCLTKFSKFSQIPAKRP 293
DB 223 QGNVKEEPRYSAKRMASEILKHPLOYEYOY-----RP 256
OY 294 ASGONSISVMPAOKITRPAKYGIPLAYKKYGDKLHEKKPLQKHKQAHQPREKRVNIGE 353
DB 257 TLSAASI-----TPREKPLNSRE 273
OY 354 ERRIISEAARRRLEFIEKEKKOKDOILSMKAEQMKROEKERLERIRINAREQGMRYVL 413
DB 274 GRRMAEESON-----SNSSSEKDMFY-----VSDKNIRYVV 304
OY 414 SAGSGEYKAPFLGSGTIAPSFSSNGOYENHAIIFDOMOORADEN-EAKMKREITGR 472
DB 305 PSNCG-----KYETEDSG-----FVDEDEDLDHVOQSAENGNIQSVSATKRPDGH 348
OY 473 GL-----PERKQGLAYERAKOYEFLOKREEMOKAKAEQHMYLARIQIRIQNTNE 527
DB 349 GILKRVHSDQRPDYIQRHPRRTIRNIMVYLKE--EKARENG----- 387
OY 528 RQOIKAKIRGEKKKPAHNSQEGSEADMRKKIESLKANA--NARAAVILEKEOLERRK 584
DB 388 -----SPRMSRNRSPSSVPTQKNNVETPSKIPKIGDI-AHSSKTNASTRIPSKL----- 436
OY 585 EAYEREKKVWEHLVAKGVASDVSPPLQHEGSGSPSKQOMRSYIVTSALKEVGVDS 644
DB 437 -----ASD-----SARTPGSFPRPHNRPVI----- 456
OY 645 LTRRETSEEMOKTNNAISSKREILRLNENLKAQDEKQMQLSDTFEIVNEDAKENE 704
DB 457 -----DSSPKLKPRNDRISP-----SPAANHE 478
OY 705 KEKSVSDRKKWEAGQGLVIRDELTLDTSTTERHTYGEVILKIGRNGSPRRAGKSPRT 764
DB 479 AEAMSVYRR-----QRPRLPRRT----- 499
OY 765 DSVLKIGLEALQOTELLENTTIRSEISPEGKYPRLITGEKKVOCISHEINPNAIYDS 824
DB 500 -----SLIHGSRQLGAD-----ISMAAKETAKLHPVSPSEEN--SHO--SRVHAS 544
OY 825 PVETKSP-----FSEASPOMSLKEGN-LEERDDLEETILOEPSGTNKE-- 869
DB 545 PVST-TPEPKRTSVGSAKGOMSESSNISLSMOAFELCDASPRYIDMTEHTTPDDHR 603
OY 870 -----SLPCTITDVIWISSEKET--KETOSADRIITIOENVS-----EDGVSTV 911
DB 604 RSCHSEVYSFPRDISSEMIRDBENHSTMRLEIPDSVSGOVQTTIHNOREEGSCPTV 663
OY 912 ---DOLSDIH-IEPGTNDQSHSKCDVYKSVQRPFFHKKVHNSPHNLVPOVOYOCSPRE 967
DB 664 LKDISPRLQSEYEPNTSOHQHD-----DKFTVKEFVSVPGRAPRLHYEP 710
OY 968 SFAPFSSHLPKPKKNNKNSLLIGLSTGLPDPANPKMLRCSLPDLSKL-----FRFLMDV 1022
DB 711 SHOVNSHDNKTIVMSQNSAL-----EKNNSHSHPRVYDVYVHVRHSFRVGSQ 762
OY 1023 PTVG--DYRQDNLE-----IDIKDE-NIKGPRSDSEDIYFEETDDLOLQASMQQLRE 1075
DB 763 PVMSVEVGVONVDMKLVNVRDEMEVRKGAAPSES-----PTTRTISSEPDSTRTERPRE 819
OY 1076 -QPEGEISEEESVLAKNSDVEPRANGTD--VADEDDNPS 1111
DB 820 PDPITNYS-ETKSFNSCSDSSPAETRTNVSFVEETPTT 857

```

RESULT 6
 G01452
 NIMA-like protein kinase 1 - human
 C:Species: Homo sapiens (man)
 C:Date: 21-Dec-1996 #sequence_rev1sion 06-Jun-1997 #text_change 24-Sep-1999
 C:Accession: G01452
 R:lin. K.P.
 submitted to the EMBL Data Library, June 1994

```

A:Reference number: G071172
A:Accession: G01452
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-445 <L0X>
A:Cross-references: EMBL:U11050; NID:9507874; PIDN:AAA19558.1; PID:9507875
C:Genetics:
A:Gene: NIK1
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
C:Keywords: ATP
F:6-271/Domain: protein kinase homology <KIN>
F:14-22/Region: protein kinase ATP-binding motif

Query Match      8.8%; Score 550; DB 2; Length 445;
Best Local Similarity 33.3%; Pred No. 1.8e-10;
Matches 145; Conservative 68; Mismatches 156; Indels 66; Gaps 13;

OY 2 EKYVRLQKIGESFGKAILVSTEDGROYVITKEINISMSKEREESRRVAVLANMKRP 61
DB 6 EDVEVLYTIGTSGSYGRQKIRKSDGKILVWKELDYSMTAEKQMLVSEVNLRLKHP 65
OY 62 NIVQYRESE--EENGSLVIMDYCEGDLFRINAQGV-----LFOBDQILDMFVOICLA 115
DB 66 NIVRYDRIIDRTNTLYIVMEYCEGDLASVIT--KGTEROYLDEFLRVMTQLTLTA 123
OY 116 LKHVHDRK-----LIHRDIKSONIFLTRDGTVOAGDFGIARVLSVTELARTCIGTPYVL 170
DB 124 LKCHRRSDGCHTVLHRLDKPANVFLDGKQVWKLGDFGLARILNHDTSFKAFTVGPTRYVM 183
OY 171 SPEICENKRYNNKSDIMWALGVLYELCTLKHAFEGASMKNNLYLKIIISGSPVSLHYSD 230
DB 184 SPQONRMVSNKSDQISWISGILYELCALMPFAFSQELAGKIRGKRRIIPRYTSD 243
OY 231 LRSLVQLFKRNPDRPSPVNSITLKEGFIARKIERFLSPOLIAEFCLTKFSKFSQPIRA 290
DB 244 LNEITIRMLNLKDYHRFSEVLEIENPLA-----DLVAD-QRNLERGRQ---L 290
OY 291 KRPASGONSISVMPAOKITRPAKYGIPLAYKKYGDKLHEKKPLQKHKQAHQPREKRVN 350
DB 291 GEPEKSDSSPVLSLKL-----KEIOLQREKALARERL 328
OY 351 TGESE--RRKISEE-AARKRRL--EFIEKKR-----OKQIISLMKAEQMR-- 392
DB 329 OKQOELCYVERLEADKLAENLILKNYSILKERKFLSLASNPELDLNPSSVIKKVYFSG 388
OY 393 QEKERLERINRAREQ 407
DB 389 ESKENIMRSENSESQ 403

```

RESULT 7
 T29771
 hypothetical protein ZC581.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_rev1sion 15-Oct-1999 #text_change 17-Mar-2000
 C:Accession: T29771
 R:Waterston, B.; Gattung, S.; Le, T.T.
 Submitted to the EMBL Data Library, May 1997
 A:Description: The sequence of C. elegans cosmid ZC581.
 A:Reference number: Z20682
 A:Accession: T29771
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-357 <WAT>
 A:Cross-references: EMBL:AF003134; PIDN:AA854139.1; GSPDB:GN00019; CESP:ZC581.1
 A:Experimental source: strain Bristol N2; clone ZC581
 C:Genetics:
 A:Gene: CESP:ZC581.1
 A:Map position: 1
 A:Introns: 31/3; 81/1; 120/3; 186/1; 226/3; 260/1; 288/3
 C:Superfamily: Ca2+/calmodulin-dependent protein kinase I; protein kinase homology

Query Match 8.7%; Score 540.5; DB 2; Length 357;
 Best Local Similarity 37.6%; Pred. No. 2.8e-10;
 Matches 121; Conservative 60; Mismatches 118; Indels 23; Gaps 7;

1 MEKYVRLQKIGSGFGKALIVKSTEDG--ROYVKEINISRMSSKREESREAVLANM 58
 Db 1 MDNYKAVVGRGAFVGCWMLCRGKNDSHQKVIKILINHGMEKEENSIOSEVWLKKV 60
 QY 59 KRPNTVOYRESFEENGSLIYMDYCEGGDLFEKRIINAOKV-----LFOEQDILDMFVQ 111
 Db 61 QHPLLIGYIDSFIMDNQGIYMOYAEGLLERLINDQRAIKDSNMREYPERKTVLDYFQ 120
 QY 112 ICLAKHYHDKKILHROIKSONIFELTKDGV-OLGDEGARVINSTVELARFCIGTPYL 170
 Db 121 ILIALMHMHOKIVHROKIPONILMNRKTYLAKISDQISKEL-GTKSAASVIGTPYL 179
 QY 171 SPEICENKPYNNKSDIHALGCVLYELCTLKHAPAGSMKLVYKIISSGFPPVLSHYD 230
 Db 180 SPECIESRPNOKSDMSGLCVLELQLEKAFDGEHUPAIWKITRSKONPLGDHVSND 239
 QY 231 LRSIVSQLEFRNRPDRSPVNSILEKGIATKIEKFLSPOLIAEFCLTKFSKGSQPIA 290
 Db 240 VKMLVNLKTRTDKRRPVSOLS-----DPLVPLIHLHDL---GRLEPPPTK 288
 QY 291 KRPASGONS-ISVMPAOKITRP 311
 Db 289 RKPSALSRLRTYPTOSTLRP 310

RESULT 8
 A43734
 probable protein kinase nima (EC 2.7.1.-) - *Emmericella nidulans*
 C:Species: *Emmericella nidulans*, *Aspergillus nidulans*
 C:Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 24-Sep-1999
 C:Accession: A43734; S28786
 R:Osmani, S.A.; Pu, R.T.; Morris, N.R.
 Cell 53; 237-244, 1988
 A:Title: Mitotic induction and maintenance by overexpression of a G2-specific gene that
 A:Reference number: A43734; MUID:88194523
 A:Accession: A43734
 A:Residue type: DNA
 A:Residues: 1-699 <OSM>
 A:Cross-references: GB:M20249; NID:g168065; PIDN:AAA3316.1; PID:g168066
 C:Genetics:
 A:Gene: nima
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
 C:Keywords: ATP; autophosphorylation; nucleus; phosphoprotein; phosphotransferase; serin
 F:9-295/Domain: protein kinase homology <KIN>
 F:17-25/Region: protein kinase ATP-binding motif

Query Match 8.5%; Score 530.5; DB 2; Length 699;
 Best Local Similarity 27.8%; Pred. No. 1e-09; Indels 139; Gaps 18;
 Matches 168; Conservative 92; Mismatches 206;

2 EKYVRLQKIGSGFGKALIVKSTEDGROYVKEINISRMSSKREESREAVLANMKH 61
 QY 9 DRYEVLKRIKIGSGFGIIRKVRKSDGFIICRKEINIKSTEREOLTAEFNILLSLRH 68
 Db 62 NIVQY--RESFEENGSLIYMDYCEGGD-----FKRIINAOKGVLPDEQDILDMFVQIC 113
 QY 69 NIVAYVHREHLKASQDILYMEYCGGDLVWTKNLRKPKVA---EEDFWRILISQYL 124
 Db 114 LALKVH-----DRKILHRDIKSONIFELTKDGVOLGDIARVANSY 146
 QY 125 TALYRCHYGTDAEYGSNLTGPAPKPSGLKKGQAOMTILHRDLPKPNIFLGSNDVYKLD 184
 Db 147 FGIAVRLNSTVELARFCIGTPYLSPEICENKPYNNKSDIHALGCVLYELCTLKHAFENG 206
 QY 185 FGISKLMSHH-DFASTYVGTPTPYMSPEICAEKTYLRSIDIAVAGCINMYELCQSEPPFNAR 243
 Db 207 SMKNLVKLIISGFPVSLHYSYDLRSIVSQLEFRNRPDRSPVNSILEKGIATKIEKFL 266

Db 244 THIQVOKIRGKRPAPLPDEYSSELKNVIAQLRVDPDRPTATLINTPYI----- 295
 QY 267 SPOLIAEFCCTKTSKGSQPIPAKRPASGONSISVMPAOKITRKPAKYGIPLAKKYGD 326
 Db 296 --RLMRREVELNINS-----RAARKREATQOKAKDVEQAFK----- 331
 QY 327 KILHEKKPLOKHOAHOPKERVNTGERRKISEEFAARRRLTEFEKKOKQDIISLMK 386
 Db 332 -----LEKEKOIRS-----ELENISIRREWEVAKRLIROYONELDKLRKFE 375
 QY 387 AQOMKR--QEKERLERINAREOGRVNLSAGSGGEVKAFLPGSGTIAPSFSRGOYE 444
 Db 376 CEYODVAQVEYDKORANNTYBEDA--SLRSSHSSQK-----SSNSSDSDPSSSTO-- 425
 QY 445 HVAIFDQOQOARADNAKKKRELIGRGIPERQKGLAVERAKOYEERLQKREAMONK 504
 Db 426 ----ISQLSLESPYNNKAK-----LPKK-SKTPPTRSKTYV-----DSPMDI 463
 QY 505 ARAEGHMYLARLROIRLONERNOQIKAKLIGKEKRAHNSBOGSESEADARRKRIEEL 564
 Db 464 QMAEPSISIALS-----SLSPKTSNAYS-GKNIFAGEKRRKRPFEPT 506
 QY 565 KAHAN 569
 Db 507 LAYS D 511

RESULT 9
 T37970
 probable G2-specific protein kinase (EC 2.7.1.-) - fission yeast (*Schizosaccharomyces*
 C:Species: *Schizosaccharomyces pombe*
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000
 C:Accession: T37970
 R:Badcock, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
 Submitted to the EMBL Data Library, September 1997
 A:Reference number: Z21758
 A:Accession: T37970
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Residue type: DNA
 A:Residues: 1-722 <RAD>
 A:Cross-references: EMBL:Z98975; PIDN:CA11553.1; GSPDB:GN00066; SPDB:SPAC19E9.02
 A:Experimental source: strain 972h-; cosmid c19E9
 C:Genetics:
 A:Gene: SPDB:SPAC19E9.02
 A:Map position: 1
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
 C:Keywords: phosphotransferase; protein kinase

Query Match 8.2%; Score 514.5; DB 2; Length 722;
 Best Local Similarity 24.2%; Pred. No. 3.2e-09; Indels 257; Gaps 28;
 Matches 199; Conservative 116; Mismatches 250;

1 MEKYVRLQKIGSGFGKALIVKSTEDGROYVKEINISRMSSKREESREAVLANMKH 60
 QY 1 DRYEVLKRIKIGSGFGIIRKVRKSDGFIICRKEINIKSTEREOLTAEFNILLSLRH 68
 Db 61 NIVQY--RESFEENGSLIYMDYCEGGD-----FKRIINAOKGVLPDEQDILDMFVQIC 113
 QY 69 NIVAYVHREHLKASQDILYMEYCGGDLVWTKNLRKPKVA---EEDFWRILISQYL 124
 Db 116 LKRVH-----DRKILHRDIKSONIFELTKDGVOLGDIARVANSY 146
 QY 120 LYRCHYGENAPACDSQMPREIFHRKOSVLIHRDIPKPNIFLGSNDVYKLD 179
 Db 157 VELARFCIGTPYLSPEICENKPYNNKSDIHALGCVLYELCTLKHAFENG 206
 QY 180 RVFTQSYVGTPTPYMSPEICAEKTYLRSIDIAVAGCINMYELCQSEPPFNAR 243
 Db 217 SGSPFPVSLHYSYDLRSIVSQLEFRNRPDRSPVNSILEKGIATKIEKFLSPOLIAEFC 276

Db 240 OGNLSCMDHHYSDVFFLLIRHCLLEVNSDLRPTTYQLLRSPILSDIRKLESEERVLEQSD 299
 QY 277 LKTFKFSOPPIPAKRPASGONSISVMPACKIRPAKAVGIPLAYKKYDDKILHEKKPQ 336
 Db 300 L-----K 303
 QY 337 KHKHONPEKRVNTGERRKISSEAKRKRLFEIEKEKKQKQIISLKAQOMKQOE 396
 Db 304 KHMQLQL-ENDDQFREQL-----SKRESELENIASR-----LAOREILRELE 349
 QY 397 RLEIRNRAREOGKRWNVLSAGSGEVKAPLGSGLTAPSSFSRGQYEHYHAIPOMOQ 456
 Db 350 KOLRMDARYO-----RHMQTVVNSMOM 373
 QY 457 R---AEDNEAKMKREIYRGRLPEROKGLAVE---RAQ-----VEEFLOK--- 496
 Db 374 RYTSFVDHNEQ-----PSSYAEHFVDCCTEASQSPLLHLPKLGISKPLQTLSC 422
 QY 497 -----KREAMONKARAG-H-----MYVLARLQIRLQNFNEROOIKAKLAGE 538
 Db 423 PGFTLTTOOPIKRPRLKELSRALHTTATLTKYRAMSSLRTPIDKGOI-TSLQOK 481
 QY 539 KKEAHNSGEGSEADMRKRIEISLKAHANARAVALKEQLEKREKREKREKVEEHL 598
 Db 482 NQTSN-----QVADCMNK-----LLHTSLDGKTLSPSELCKNFSD--- 516
 QY 599 VAKGVSSDVSPPLGONHETGSPSKOOMRSVISTALKFVGVDSSLTPRETSE----- 652
 Db 517 -GEGLPNRKVS-----KLSVESDEFAVASAGS-ESVPTDSTLTDTKSKSVYVHP 564
 QY 653 -----EEMOKTN---NAISSKREILRLN-----ENLKAQDEDEKGMON 687
 Db 565 SPOSILYKLEKLIIRSDVEKSPKASKTILGALPSLASPIYVAHEKTIARNE--MDG 622
 QY 688 LSDTFEIVHEDAKHEKEKSVSDRKKEWAGGOLVIPLEDEL 729
 Db 623 NFKTKINQHPD-----EYVLRTPKRIQLLEQOKSPVKQL 658

 RESULT 10
 S59359
 N:Alternate names: yeast (Saccharomyces cerevisiae)
 C:Species: Saccharomyces cerevisiae
 C:Date: 30-Nov-1995 #sequence_revision 01-Mar-1996 #text_change 24-Sep-1999
 C:Accession: S59359; S69565
 R:Longline, M.S.: Pringle, J.R.
 submitted to the EMBL Data Library, August 1995
 A:Reference number: S59359
 A:Accession: S59359
 A:Molecule type: DNA
 A:Residues: 1-1142 <ION>
 A:Cross-references: EMBL:U03140; NID:g992650; PIDN:AAA7513.1; PID:g992651
 R:Dieckrich, F.S.
 submitted to the EMBL Data Library, August 1995
 A:Description: The sequence of S. cerevisiae cosmid 8166, 9787, 9717, and lambda 3073.
 A:Reference number: S69565
 A:Accession: S69565
 A:Molecule type: DNA
 A:Residues: 1-1142 <DIE>
 A:Cross-references: EMBL:U03057; NID:g927764; PIDN:AAB64949.1; PID:g927777; MIPS:YDR507C
 C:Genetics:
 A:Gene: SGD:GIN4
 A:Cross-references: SGD:S0002915; MIPS:YDR507C
 A:Map position: 4R
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol
 C:Keywords: ATP
 F:17-289/Domain: protein kinase homology <KIN>
 F:25-33/Region: protein kinase ATP-binding motif

Query Match 8.0%; Score 502.5; DB 2; Length 1142;
 Best Local Similarity 19.8%; Pred. No. 1.1e-08;

Matches 252; Conservative 215; Mismatches 475; Indels 331; Gaps 42;
 QY 8 OKIGESFGALIVKSTEDGROYVKEI-----NLSRMSKREES-----REV 52
 Db 23 ETGLGSLTGQVLARNGSTGGEAAVKISKAVNFNGVNSGTSYGSTTPPALPGIEREI 82
 QY 53 AVLANKHRIYVYKRSFENSITYVNDYCGGLDFRINAQGVLPQEQILDPVQI 112
 Db 83 IIMKLHPVNLVLYWETNDLTVLETAAGGLFLL--VEKGLP-PHEAARFRPQI 140
 QY 113 CLAKHVHDKILHRDIKSONIFLTKDQVGLDGFARLVNSTVELARPGIGTPYVLS 172
 Db 141 IIGVSYCHALGIYHRDLKPENLILDRKYNIKIADGMA-ALETGKLETSCGSPHAP 199
 QY 173 EIGENRPYNN-KSDIWLGCVLVELCTLKHAFF--EAGSMKVLVLTISGSP-PVSLHYS 228
 Db 200 EIVSGIPIYOGFASDVWSCVILFALLTGLRPDEEDGNIRLLLVOKGFEEMPSDEIS 259
 QY 229 YDLRSVLSQLFKNRPDRSVNSILEKGFIAK-----RIEFLSPQILAE 274
 Db 260 READDLIRKILVDPERRIKTRDILKHPLOKYPISRDSKIRGLPREDDYVLPFSNS 319
 QY 275 FCLKTFESK-----FGSOPPIPAKRPASGONSISVMPACKIRPAK-----YGIPLAYK 323
 Db 320 SIDATILQNLVILMHGRDPEGIK-----EKLEPGANMEKTLVALLYRFK 365
 QY 324 YGDKILHEKKPILOKHQAQO---TPEKRYNTGEERR-----ISEPARRKRLFEIEK 375
 Db 366 DTQKELIKQOQVKKRQSSISVSFSPKYSTTPQRRNRRESLIVTSRKRPISF-NKFT 424
 QY 376 KQKQIISLKAQOMKQOEKRELEIRNAREOGKRWNVLSAGSGEVKAPLGSGLTAP 435
 Db 425 ASSANSSNLTPESSKR-----LSKNFSSKKLSTIVMSSPTPA 464
 QY 436 SFSSROYEYHAIPOMOQOAEDENAKMKREIYGRLP-----EROKGLAVE 485
 Db 465 SRNRKAS---VINENKOKRASIFSTYKRRSSRSIKRMSLIPSKRRSVYTKMT 519
 QY 486 RAKQVE---EFLQK-REAMONKARAGHMYVLARLQIRLQNFNEROOIKAKLEKK 540
 Db 520 YAKLAEDDWEYIEKTKRTSSNFATLIDEIFEYKYQIR---KKEELERKVRBAK- 574
 QY 541 EAHNSGEGSEADMRKRIEISLKAHANARAVALKEQLEKREKREKREKVEEHLVA 600
 Db 575 -----ARELEERRRQOEKE---RAKRLKEKDLKQOEL---KQIE----- 613
 QY 601 KGVSSDVSPPLGONHETGSPSKOOMRSVISTALKFVGVDSSLTPRETSEEMOKTN 660
 Db 614 --IDIDLQEGELSKH-----KEKIDGNIRISAPMEHEKININ 650
 QY 661 AISKRE-ILRLNENLKAQDEKGMONLSDFEIVHEDAKHEKEKSVSDRKKEAG 719
 Db 651 HLEVYDNLIRKRNFSLOIR-----KEKIDGNIRISAPMEHEKININ 650
 QY 720 GQIVLPDELTLDTSSPTERTVGEVYIKLGPNGSPRRAMKSPDVSYLKIGEAELQLO 779
 Db 671 ----PVSRDPCIMFS-----SPEEEV----- 688
 QY 780 TELLENITIRSEISPEGEYKPLITGEKKVOCI--SHEINSAIVDSVPETKSPSEAS 837
 Db 689 -----SPEPKRTENERLITTEKILLETIRSKVLGSSFNIDELKLSKMEYPSII 738
 QY 838 PQMSLKLEGNLEPPDLEIILQEPSTGNKDESLCTTIDVWIE----- 882
 Db 739 APQRLSEERVSDNSNGYESLILPDKNGVSLKQSTATAPVSDGLRKISERIVPQFT 798
 QY 883 EKETKFGQSDRTTQIENFVSDGVSVYDOLSDIHPGNDGSHKCOVDKSVQDEP 942
 Db 799 RKRHRFSESUKRSLVLSMTSKESFTNLVDILKQNDL--VNNQSGRIRTPRSADSE 856
 QY 943 FKHVYSEHLNVPOVQSCPESSPAFRSHSLPFRNKKNSLLGLSTGLFDANPK 1002
 Db 857 LFEIVNE-----AETGNSNDRRLVDVGDSITIKDKSLKLNLFADRFNGSNEAK 906

QY 1003 MLRTCSLPDLSKLERTIMDVPGVQKDNLEIDEIKDENIKEGSS---DSEDIYFEETD 1059
 DB 907 QTDNHLILP-----PLNDG-----NELRKNOSGOGAHPKISMLPEGS 949
 QY 1060 TLOELQASMBOLLREOGEYESEESVLKN---SDVEPTANGTDVADDDNPSSAL 1116
 DB 950 SSHTEKEENEKEEKKP-EQHKOEEOEKREKVDMEPLKSVQIKREKNAQSA-- 1006
 QY 1117 NEMHSDNSDGEIASECDSVFNHLEELRLHLEOEWKPEPEYEYKIKALHEDEENI 1176
 DB 1007 -----KHSK-----DHLKEHKODKNTAIGNSFFRKFSK-----SSDKTM 1042
 QY 1177 EICSKTYONILGN 1189
 DB 1043 ELTKAKISAKOLF 1055

RESULT 11
 T14157
 serine/threonine protein kinase - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T14157
 R:Pytowski, B.; Hicklin, D.J.; Kornhaber, G.; Dellaratta, D.V.; Witte, L.
 submitted to the EMBL Data Library, December 1997
 A:Reference number: 217894
 A:Accession: T14157
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1233 <P>Y
 A:Cross-references: EMBL:AF039574; NID:g2773155; PID:g2773156; PIDN:AA896682.1

Query Match 8.0%; Score 500; DB 2; Length 1233;
 Best Local Similarity 20.8%; Pred. No. 1.5e-08;
 Matches 272; Conservative 224; Mismatches 508; Indels 304; Gaps 49;

QY 9 KIGSGFGKALIVSTFGROYVYKEINISRMSSKERESEREAVANLNMKNPIVQYRE 68
 DB 39 ELGGAGGKGYKAKQNKETNVLAAKVIDRK--SEELDYWEIDILASCHPPIVILKD 96
 QY 69 SPEENGSLXTMDYCEGGDLFKRINAKQGVLFQ-----EQDILDMFOICLALKHVHR 122
 DB 97 AFYENNMILIEFCAG-----AVDAMLELERPLEESQIOVVCQOTLEALNYLHDN 149
 QY 123 KILHDIKSONIFLTGQVQLODPIGIAVNLSTVELARTCIQTPYVLSPEI--CE--N 177
 DB 150 KLIHDLKAGNIFLTLDGDLKADGVSAKNTIQRDSFICITPYWMAPEVMCETSKD 209
 QY 178 KPYNKRSDIMALGVLYELCTLKAHFAFGSMKULYKIIISGSPRPS--LHYSYDLRLV 235
 DB 210 RPYDYKAVWMSIGITLLEMAEIEPPHMLPMKRLIKASEPTILAQSKMSNFKDPL 269
 QY 236 SOLFRNRPDRPSVNSILEKGFIAKRIEFLSPOLIAEFCUKTFS--KFGSQ----- 286
 DB 270 RKLCEKNDVARWTTSQLDHPFVVDNSKRPV-ELIAEAKAEVTEVEGKEDEBEARE 328
 QY 287 ---PIAKRPASGONSISVMPAQKITRPA-----KYGIPLAYKKYCDKILHEKKPLQ 336
 DB 329 NALPIPNKRASSDLSIASSEEDKISQNAKLESVSRTOSTSEKFSKILINKEPTTD 388
 QY 337 KKHQAHQPEKRYN--TGEERRK-----ISEFAARRLEFIEREKQKQDIISL----- 384
 DB 389 GPKAVDEHNASDVNLETGALNDQVGIHNGRKKRPKLENLPDODQDVVNSVSEE 448
 QY 385 -----MKAPMKROKE-----RLERINRAREOGNRNVLASGSGEYK 422
 DB 449 NENNRVLTETNDCKLPEDRKNEQETLESKLIQSEIINTHIOYMD--LVSOETGEKE 506
 QY 423 APF-----LGGGTIAPSSFSRQGEYENHAIFFDQMOO-----OR 457
 DB 507 ADFOAVDNEVGLTKEETQELKDGTAQKVITSRRSSVPTGDEALDQAKAELSKRAAS 566

QY 458 AEDNEAKKREIYRGILPEROKQLAVERAKOYEEFLQKKREAMONKARAGHWYLARL 517
 DB 567 GEGDEALVPQTQIAKEPTEGPEAGGAEEPPGGERVEDQAPQOAVCEALGOLTSSET 626
 QY 518 RQIRLO--NFEROOI-----KAKLGEKKAHNSBEOGSEADMRKKIES-- 563
 DB 627 TQATLEQETDEVRQVSESNIELERLVYTGAEARALISEEAATEDELRKKNQKV 686
 QY 564 -LKAAHNRAA-----VLKEOLEKRRKREARERKKWEHILVAKGVSSDVSPLQ 614
 DB 687 PKAKSQAPASQPSPPHVLIPISINSETTENKE-----MALPKPETILPEPE 739
 QY 615 HETGG-----SPSKOQMSVTSALKEVGSDSLDRTRESEEMQKTNAI- 662
 DB 740 HEKNDTSGTGTIVENSGLNLSSPLSKAKDSG--SVLSQETRQKTLKTRKIV 798
 QY 663 -----SSKRLIRLME-----NLKAOEDEKGMNLDSTFEINVHEDA 700
 DB 799 DGEVSVTTSKIVTSDSKTELEFLRLQRLRELRLQKEGRQAQOLNGKIQOOREQIF 858
 QY 701 KEHEKESVSDRKWEAGQOLVPIDELTDTSFTTERHTYGEVYIKIGPNSPRAMG 760
 DB 859 RREOE--MLSKKROY--QETENLEKQOKOTIERLEQHT----- 895
 QY 761 KSPDLYKILGAELOLO--TELEEN--TIRSEIISPBGEKKYKPLITGKKYQICSHI 816
 DB 896 NRLRDEAKRIKGOEKELSKFQNVLNKKKEVMNDEKARRELRLKTRKRELQSQ- 954
 QY 817 NPSAIVDPVETKSPFSEASPOMSIKLEGNL-----EPDLETETLOPSTGND 868
 DB 955 -----HAQOEFGVQKQO--ELDGLTKIIQOKAELANITREL-----NNQ 996
 QY 869 ESLPCTITDVWISSEKETE-----TQSDRITIGNEVSDGVSTVQSLDIHIEG 922
 DB 997 QLMARARALWELEHRLDEKHLQKQKLDQYFMORHQLK----- 1038
 QY 923 TNDQSHKCVQDSVQPEPFHVV-----HSEHLNVPOVQSVQSPESF--AFNS 973
 DB 1039 -----RHKK-----ETEQMQRNRLIELEKNTQTORARLPKIQSEKTRAMFKSLRI 1090
 QY 974 HSHLP---KNKNNSLLIGLSTGLDANNPMYLTGSLPLSKLFTLMD--VPTVGVY 1028
 DB 1091 NSTATPQDRKRIKO-----FPAQDEKRNKERRAAGHOKHESQMRDLOLCERANY 1140
 QY 1029 RODNLEIDEIKDENIKRSPDSSEDIYFEETDDLOELQASMBOLLRE-----QGEYESE 1083
 DB 1141 R---ELHQLONEKC-----HLYVHETQKTLKLEDEHGOELKEVREKILRRKRTLE 1188
 QY 1084 EE-ESVLKNSDVEPTANGTVADEDDNPSSSALNEM-----HSDNS 1125
 DB 1189 EEFARKLOEVEFFRMKG---ESECLNPSAKQSKISFYIPIPLHSTGS 1233

RESULT 12
 A57177
 NIMA-like protein kinase - Emericella nidulans
 C:Species: Emericella nidulans; Aspergillus nidulans
 C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 24-Sep-1999
 C:Accession: A57177
 R:Pu, R.T.; Xu, G.; Wu, L.; Viterula, J.; O'Donnell, K.; Ye, X.S.; Osmanli, S.A.
 J. Biol. Chem. 270, 18110-18116, 1995
 A:Title: Isolation of a functional homolog of the cell cycle-specific NIMA protein ki
 A:Reference number: A57177; MUID:95355415
 A:Accession: A57177
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-779 <P>UA>
 A:Cross-references: GB:I42573; NID:g1040682; PIDN:AA80145.1; PID:g1040683
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
 C:Keywords: ATP
 F:3-290/Domain: protein kinase homology <Kin>
 F:13-21/Region: protein kinase ATP-binding motif

Query Match 8.0%: Score 499.5; DB 2; Length 779;
 Best Local Similarity 22.0%: Pred. No. 9.9e-09;
 Matches 211; Conservative 129; Mismatches 326; Indels 291; Gaps 31;

2 EKYVRLQIGGSGFKALVSTEDGRQYIKENINISMSSKREESREAVLANMKHP 61
 5 DKYLELEKIGHGSGFIIKRRARRADGMLCKREISYLKSKOKEREDLHAFESILSTLHP 64
 62 NIVQY--RESPENGSLIYIMVDCGEGDFKRI--NAQKGVLFQEDQILDW--FVOICLAL 116
 65 NIVGYHREHLKATQDILHLYMEYCGNDLGRVIRNLKNNQVAEESEFV--WSIFQSLVAL 123
 117 KHVH-----DRKILHRDIKSONIFLTKDGTVOLGDFGIA 150
 124 YRCHGVDPPEVKGTVLGLSTARPKPPSGCMTILHRDLKPEVFLGEDNSYKLGDFGLS 183
 151 RVLNVSTELARTCIPTFYVLSPEICENKPPYNNKSDIMALGCVLYELCTLKHAFEGSMKN 210
 184 KVMQSH-DFASTYVGTPTFYSPELCAEKTILKSDIMSGLCITIELCAHEPFPNKATHYQ 242
 211 LVKATISGFPVSLHSTYDLRLSVLSQLEFRNRPDRPSVNSLIEKGFIAKRIEKLSPOL 270
 243 LVQKREKRIAPLPSVLSGELFATIKDCLRVNDRPDRTATL----- 285
 271 IAEFCILKTFSGSQPIPAKRPASGONSISVMPAOKITKPAKYGIPLAYKYGDKKLH 330
 286 -----NLPTV-----RLMR 294
 331 EKKPLQKHQAHTQPERKRVNTGERRKISEEAAKRRLEFIEKKKKOKDOITSLMKADOM 390
 295 KEKEVEFSRFLRTKEETLN---KRIRE--LDKSLALETEK-----SSTRAE-- 337
 391 KROEKERLERINRAEOGMNRVLVAGSGGEVKAPFLGSGGTIAPSSFGQYEHYHAF 450
 338 -----IDASLRREM-----EVKA-----RLKI 354
 451 DOMQOQRAEDNEAKMKREI-----YGRG--LPEROGGOLAVAKOVEEFLQK 497
 355 DRLVQLESLQKFEQEOQARVAELQHRGRPMNSHGQCSFSTATILVSDY----- 410
 498 REAMONKRAEGHNVYLAIRLOIRLQNFNRQOIKAKL---REKKEAHHSEGOBS 551
 411 --NLSSVSGSGDPPSTTDITDITIAESTGSDITKIPRTPPHRAQITSSAPASVLAGT 468
 552 ----EADMRKKITESLKAHANAAVAKLEQLEKRR---KEAYEREKRVWEHLVANGV 603
 469 PMDIEMASPPITIASLST--SPRMALTKAPTNPIMFGEETSTDKSNMVEVPRETEMI 527
 604 KSSDVS-----PPLGQHEHG-----GSPSKOQMSV 629
 538 DSGGESEALVPSPKRITKSKNPFSTVTRRPSLNSQOANSVLPJIGLSKQOTLAIR 587
 630 ISVTLSALKEVGD-----SLTDRTRESFEMOKTNNALISKRILRLNENKAOEDERG 684
 588 SKTVSGVSGISQHPRLRSAPSLRDKRPSRRLSRIPSVTCVGRRLANNINNSNGG 647
 685 MONTSDPFEIIVHEDAKHEHEKESVSSDRKKMEAGGOIYLPDELDTLQTSFSTERHYG 744
 648 SDAPSTYTSMTIVTRGKLRMSSTCD-----ESSFQOO----- 682
 745 EYIKIGPGRSGRRAMGSPDSDVYKILGEALQLOTELENTTIREISIP-----GE 797
 683 -----NNQPOSLPQAPP---LTKTG-----LMAAKNIRSSSLVELHQAARAGR 723
 798 KYKPLITEKKVQCI--SHEINPSAIVDSPVETKSPESSEASPMSLKLGNELEPPD 853
 724 PISATISNEAKLRAFKEHATIASAVDS-----SSSSSSSSGOSQLPTRPRSQPPTD 775

RESULT 13
 B96587

hypothetical protein F20D21.33 (imported) - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: B96587
 R:Thelodopsis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
 anseu, N.F.; Hughes, B.; Huizart, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marzla
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallo
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: B96587
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-200 <SNQ>
 A:Cross-references: GB:AE005173; NID:94585993; PIDN:AAD25629.1; GSPDB:GN00141
 C:Genetics:
 A:gene: F20D21.33
 A:Map position: 1

Query Match 7.6%: Score 475; DB 2; Length 200;
 Best Local Similarity 43.8%: Pred. No. 1.6e-08;
 Matches 85; Conservative 52; Mismatches 55; Indels 2; Gaps 2;

1 MEKYVRLQIGGSGFKALVSTEDGRQYIKENINISMSSKREESREAVLANMKH 60
 1 MEQYEFLEQIGGSGFSLVLRHKKKKYLLKRIARLQTORRRSHQEDVLISKMH 60
 61 PNIVQYRESFENGSLIYIMVDCGEGDFKRIQAQGVLFQEDQILDWVQICLAKHV 119
 61 PTVEYKDSWVKACVYCVIGYCEGDMQAQIKKSNQVHFEKELCKWLVQILMLELYL 120
 120 HDRKILHRDIKSONIFLTKDGTVOLGDFGIRAVNSVTELARTCIPTFYVLSPEICENP 179
 121 HSNHILHRDYKSNIFLTKEDDIRGDFGLKILTSO-DLTSVVGTPSMCELLADLP 179
 180 YNKSIDIALGCVL 193
 180 YGSKSDIMSLGTFI 193

RESULT 14
 S49313
 protein kinase - slime mold (Dictyostelium discoideum)
 C:Species: Dictyostelium discoideum
 C:Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 24-Sep-1999
 C:Accession: S52076; S49313
 R:Wetterauer, B.W.; Hamker, U.; von Haeseler, A.; MacWilliams, H.R.; Simon, M.N.; Ver
 Blochum, BioPhys. Acta 1265, 97-101, 1995
 A:Title: A protein kinase from Dictyostelium discoideum with an unusual acidic repeat
 A:Reference number: S52076; MUID:95161460
 A:Accession: S52076
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1094 <WE2>
 A:Cross-references: EMBL:Z37981; NID:9551445; PIDN:CAAB605.1; PID:9551446
 C:Genetics:
 A:Insertions: 35/3; 104/1; 166/2
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
 C:Keywords: ATP
 F:20-281/Domain: protein kinase homology <KIN>
 F:28-36/Region: protein kinase ATP-binding motif

Query Match 7.5%: Score 467; DB 2; Length 1094;
 Best Local Similarity 21.5%: Pred. No. 1.3e-07;
 Matches 252; Conservative 208; Mismatches 434; Indels 280; Gaps 50;

Wed May 15 14:14:04 2002

us-09-783-320-4.rpt

Page 10

```
Db 513 -AIDNEVGFTKEQKIGKDKDKTKHV-----ISDI 543
QY 633 TSALKVEGVDSLTDTRETSEMOKTNNAISKR-ETLRRLNENLK-----AOEDEK 684
Db 544 TS---EVGTDERPGDTQKSAEQSDABEGAGEAPQAQITTEKATEGPEAHGAEERPS 600
QY 685 MONTSTFEINNHEDAKHEHEKESYSDRKWEAGOLV-----IPDELTLDTSFST 737
Db 601 GERVED-----KOEQOSAVC-----EGEQVTSSESTRATTEPETEDVOY 644
QY 738 TERHIVGEYIKIGPWSPPRANGKSPDTSVLKIIGEAELQLOTELLNTTIRSEISPEGE 797
Db 645 SESNSTEELERLGVTAEBQALGSK-----GEATELDERENA---OELPVKAE 692
QY 798 KYRPLITGKKVOCISHEINPSAIVDPETKSPFESEASPOMSLKEGNLEEPDLETE 857
Db 693 POAP-----AASQASEPPVLIPIINIH-ENTENKGE-----GALPRP---ETI 734
QY 858 ILQEP---SGTNKDESLPCTI---TDVWISSEKETETOSADRIITIOENVESEDGVSSP 910
Db 735 LPPEPENGKNDTSGTSTVENSSDLNISISFLSKTKDGSVSLQETRRQKTKTKT 794
QY 911 VPOQLD-IIEPGTNDOSKCDVKSVOPEPFHKVYHSEHLNLPVOYOSVOCSPESF 969
Db 795 RKIVDGVESVYT-----SKIVTDSKTEEL--RFLRQELRELRLLOKEGKAQQL 847
QY 970 AFRSHSLPPKKNKNSLLIGISTGLFDANNPKMLRTCSLPDLSKLFRILMDVPTVGDVR 1029
Db 848 -----NGKLOQOREQIFRFEQEMLSKKRQYDOELENL--EK 882
QY 1030 QDNLEIDETKDNKEGSDSEDIYFEETDTLOELQASMEQLRQPGEE-----YSE 1083
Db 883 QOKQTERLEQEHNLRLDEARIRKEQ-EKELSKFO---NMLRNRKKEQEFVOKQOQ 937
QY 1084 EESVVLKNSDVEPTANGDTVADEDDNPSSSESALNEE--WHSNDSGEIASCECDSVFN 1140
Db 938 ELDGALRKLIIQOKKALANIIEKCLNNKQOLLRARAAIWELE-----R 982
QY 1141 HLEELRLHLEQEMGEFEVEYEKIKAIHEDDENIE 1177
Db 983 HLOEKHQLKQOLKDQYFIORHQLKR-HEKETEQMOQ 1018
```

Search completed: May 15, 2002, 07:55:09
Job time: 219 sec

Wed May 15 14:14:05 2002

us-09-783-320-4.rsp

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 15, 2002, 07:57:59 ; Search time 33.64 Seconds
(without alignments)
1397.311 Million cell updates/sec

Title: US-09-783-320-4

Sequence: 6243
1 MEKRYVLRKIGSGFGRKIL.....YAKILHLVMDAGNYEDNDE 1214

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SWISSProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	3385	54.2	774	1 NEK1_MOUSE
2	847.5	13.6	841	1 SMP2_HUMAN
3	833	13.3	511	1 NEK3_MOUSE
4	790.5	12.7	459	1 NEK3_HUMAN
5	565.5	9.1	431	1 NEK4_TRYPB
6	550	8.8	443	1 NEK2_HUMAN
7	544.5	8.7	443	1 NEK2_MOUSE
8	530.5	8.5	699	1 NIMA_KENT
9	528.5	8.5	431	1 NIMA_TRYPB
10	502.5	8.0	1142	1 GIN1_YEAST
11	499.5	8.0	779	1 NIMA_NEUCR
12	468.5	7.5	966	1 ST10_MOUSE
13	455	7.3	966	1 ST10_HUMAN
14	452	7.2	1518	1 KKK1_YEAST
15	437.5	7.0	915	1 KCC4_YEAST
16	432	6.9	435	1 KIN3_YEAST
17	421.5	6.8	1895	1 YLK3_CAEEL
18	393	6.3	1080	1 NRK1_YEAST
19	388	6.2	1062	1 CC7_YEAST
20	383.5	6.1	705	1 CC5_YEAST
21	382.5	6.1	740	1 K6A3_HUMAN
22	382	6.1	713	1 KP78_HUMAN
23	381	6.1	460	1 CDS1_SCHPO
24	380.5	6.1	685	1 SNK_HUMAN
25	380	6.1	682	1 SNK_MOUSE
26	378	6.1	631	1 SNK_MOUSE
27	378	6.1	733	1 K6A2_HUMAN
28	378	6.1	733	1 K6A2_MOUSE
29	378	6.1	735	1 K6A1_HUMAN
30	378	6.1	735	1 K6A1_MOUSE
31	377.5	6.0	733	1 K6A1_RAT
32	377.5	6.0	1305	1 GAK_RAT
33	373	6.0	1	1 GAK_RAT

34	374.5	6.0	982	1 SULK_CAEEL	P46549 caenorhabd
35	374	6.0	350	1 KAP6_HUMAN	P22612 homo sapien
36	374	6.0	615	1 CNK_RAT	O97011 rattus norv
37	374	6.0	745	1 K6A6_HUMAN	O9AK32 homo sapien
38	373.5	6.0	490	1 SP51_YEAST	P08458 saccharomyc
39	372	6.0	646	1 CNK_HUMAN	O914B4 homo sapien
40	371.5	6.0	485	1 K6B2_MOUSE	O921M4 mus musculu
41	371	5.9	350	1 KAP4_BOVIN	P00517 bos taurus
42	371	5.9	350	1 KAP4_HUMAN	P17612 homo sapien
43	371	5.9	774	1 KEMK_MOUSE	O05512 mus musculu
44	370.5	5.9	444	1 KRAC_DICDI	P54644 dictyosteli
45	369.5	5.9	397	1 KAP4_YEAST	P06244 saccharomyc

ALIGNMENTS

RESULT	ID	STANDARD	PRT	774 AA.
1	NEK1_MOUSE			
AC	P51954	01-OCT-1996 (Rel. 34, Created)		
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Serine/threonine-protein kinase NEK1 (EC 2.7.1.-) (Nima-related protein kinase 1).			
DE	NEK1.			
GN	Mus musculus (Mouse).			
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Blood;			
RA	Medline=93010942; PubMed=1382974;			
RA	Letwin K., Mizzen L., Motro B., Ben-David Y., Bernstein A.,			
PT	NIMA cell cycle regulator and highly expressed in meiotic germ			
RT	cells".			
RU	EMBL J. 11:3521-3531(1992).			
CC	-1- FUNCTION: PHOSPHORYLATES SERINES AND THRONINES, BUT ALSO APPEARS			
CC	-1- TO POSSESS TYROSINE KINASE ACTIVITY. IMPLICATED IN THE CONTROL OF			
CC	MEIOSIS.			
CC	-1- SUBCELLULAR LOCATION: Nuclear (Probable).			
CC	-1- TISSUE SPECIFICITY: PREDOMINANTLY IN TESTES (GERM CELLS AND			
CC	SERTOLI CELLS). LOWER LEVELS IN OVARY (OVOCYTES AND GRANULOSA			
CC	CELLS), THYMUS AND LUNG.			
CC	-1- DEVELOPMENTAL STAGE: IN FEMALE, EXPRESSED AS FOLLICLES ENTER THE			
CC	SECONDARY STAGE UNTIL OVULATION OCCURS. IN THE MALE REPRODUCTIVE			
CC	SYSTEM, THE EXPRESSION IS LIMITED TO SPERMATOCYTES AND SPERMATIDS.			
CC	-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.			
CC	NIMA SUBFAMILY.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement. (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL: S45828; AAB23529.1; -			
DR	HSSP: P24941; 1BUH.			
DR	MGD: MGI:97303; Nek1.			
DR	InterPro: IPR000719; Euk_pkinase.			
DR	InterPro: IPR002290; Ser_thr_pkinase.			
DR	Pfam: PF00069; pkinase; 1.			
DR	SMART: SM00220; S_TKC; 1.			
DR	PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.			
DR	PROSITE: PS00108; PROTEIN_KINASE_ST; 1.			
DR	PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.			
DR	transferase; Serine/threonine-protein kinase; ATP-binding; Mitosis;			
KN				

KW Nuclear protein; Phosphorylation; Cell cycle; Cell division;
 KW Tyrosine-protein kinase.
 FT DOMAIN 4 258 PROTEIN KINASE.
 FT NP_BIND 18 ATP (BY SIMILARITY).
 FT BINDING 33 33 ATP (BY SIMILARITY).
 FT ACT_SITE 128 128 BY SIMILARITY.
 SQ SEQUENCE 774 AA; 88427 MW; 7EDES81ACCC06FA CRC64;

Query Match 54.2%; Score 3385; DB 1; Length 774;
 Best Local Similarity 85.5%; Pred. No. 1, 2e-118;
 Matches 665; Conservative 45; Mismatches 64; Indels 4; Gaps 3;

QY 1 MEKYVRLQKIGSGFGKALIVKSTEDGRQYIKETINSRMSKREESREVAVALNMKH 60
 DB 1 MEKYVRLQKIGSGFGKALIVKSTEDGRQYIKETINSRMSKREESREVAVALNMKH 60
 QY 61 PNIVQYRESFEENGSLVYIMDYCEGDLFRKINAKGVLFQEDQIILDMFVQICLALKVH 120
 DB 61 PNIVQYRESFEENGSLVYIMDYCEGDLFRKINAKGVLFQEDQIILDMFVQICLALKVH 120
 QY 121 DRKILHROIKSONIFLTGDTGVQGLDFGIARVLSVTELARTCIGTPYILSPICENKPY 180
 DB 121 DRKILHROIKSONIFLTGDTGVQGLDFGIARVLSVTELARTCIGTPYILSPICENKPY 180
 QY 181 NKSQDMALGCVLYELCTLKAHAFAGSMKNLVKTIISGSPVSLHVSVDLSVQLFK 240
 DB 181 NKSQDMALGCVLYELCTLKAHAFAGSMKNLVKTIISGSPVSLHVSVDLSVQLFK 240
 QY 241 RNPDRPVSNTLLEKFLAKRIEFLSPQIAEFCLTKTSKGSQPIPAKRPASQNSI 300
 DB 241 RNPDRPVSNTLLEKFLAKRIEFLSPQIAEFCLTKTSKGSQPIPAKRPASQNSI 300
 QY 301 SVPAOKITRPAKYGIPLAYKKYGDKKLHEKRPLOKHKNOHOTPEKRVNTGEERRISE 360
 DB 301 SVPAOKITRPAKYGIPLAYKKYGDKKLHEKRPLOKHKNOHOTPEKRVNTGEERRISE 360
 QY 361 EAKRKRLFEIKERKKOKQDITSLMAKQOMKROKERYLERINBARQGMNVLSAGSGE 420
 DB 361 EAKRKRLFEIKERKKOKQDITSLMAKQOMKROKERYLERINBARQGMNVLSAGSGE 420
 QY 421 VNAFELSGGTAPSSFSRGOYEYHAIIPOMOQOAEENEMKREITGRGLPRRQNG 480
 DB 421 VNAFELSGGTAPSSFSRGOYEYHAIIPOMOQOAEENEMKREITGRGLPRRQNG 480
 QY 481 OLAVPRAKOVFEELQKREAMONKARAGHMYLARIRQIRLQNEPROQIRAKLECEK 540
 DB 481 OLAVPRAKOVFEELQKREAMONKARAGHMYLARIRQIRLQNEPROQIRAKLECEK 540
 QY 541 EANHSGGSEADMRKKITJESLKAHANARAVALKQDLERKREKAVEREKVMEEHVA 600
 DB 541 EANHSGGSEADMRKKITJESLKAHANARAVALKQDLERKREKAVEREKVMEEHVA 600
 QY 601 KGVKSDVSPPLGQHETGSGSPSKOQMSVISTSALEKGVSDSLDTRETSEEMOKTNN 660
 DB 601 KGVKSDVSPPLGQHETGSGSPSKOQMSVISTSALEKGVSDSLDTRETSEEMOKTNN 660
 QY 661 AISSKREILRLNENMLKAOEDEKQMONLDPFEINVEDAKHEHREKESVSDSKKWEAG 720
 DB 661 AISSKREILRLNENMLKAOEDEKQMONLDPFEINVEDAKHEHREKESVSDSKKWEAG 720
 QY 721 QLVIPRLDELTLTSTSTHRTYGVYIKLPGNSPRAMKSPRDSVTKITGAEIOL 778
 DB 721 QLVIPRLDELTLTSTSTHRTYGVYIKLPGNSPRAMKSPRDSVTKITGAEIOL 778
 QY 778 QLVIPRLDELTLTSTSTHRTYGVYIKLPGNSPRAMKSPRDSVTKITGAEIOL 778
 DB 778 QLVIPRLDELTLTSTSTHRTYGVYIKLPGNSPRAMKSPRDSVTKITGAEIOL 778

RESULT 2
 STR2_HUMAN
 ID STR2_HUMAN STANDARD: PRT; 841 AA.
 AC P51957;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Serine/threonine protein kinase 2 (EC 2.7.1.37) (Serine/threonine-
 DE protein kinase NRK2).
 GN STR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_Taxid-9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-PREST:
 RX MEDLINE-9426838; PubMed-8208544;
 RA Levedakou E.N., He M., Baptist E.W., Craven R.J., Cance W.G.,
 RA Welch F.L., Simmons A., Naylor S.L., Leach R.J., Lewis T.B.,
 RA Bowcock A., Liu E.T.;
 RT Two novel human serine/threonine kinases with homologies to the cell
 RT cycle regulating Xenopus Mo15, and NIMA kinases: cloning and
 RT characterization of their expression pattern.";
 RL Oncogene 9:1977-1988(1994).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphoprotein.
 CC -1- SUBCELLULAR LOCATION: Nuclear (probable).
 CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN ADULT HEART, FOLLOWED BY
 CC PANCREAS, SKELETAL MUSCLE, BRAIN, LIVER, KIDNEY, LUNG AND
 CC PLACENTA. PRESENT IN MOST PRIMARY CARCINOMAS.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC NIMA SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).
 CC -----
 DR EMBL: L20321; AAA36558.1; -
 DR HSSP: P11362; 1FGK.
 DR MIM: 601959; -
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR PROSITE: PS00107; PROTEIN KINASE. ATP. 1.
 DR PROSITE: PS00108; PROTEIN KINASE. ST. 1.
 DR PROSITE: PS50011; PROTEIN KINASE. DOM. 1.
 KW Transferrase; Serine/threonine-protein kinase; ATP-binding; Mitosis;
 KW Nuclear protein; Phosphorylation
 FT DOMAIN 6 261 PROTEIN KINASE.
 FT NP_BIND 12 20 ATP (BY SIMILARITY).
 FT BINDING 35 35 ATP (BY SIMILARITY).
 FT ACT_SITE 131 131 BY SIMILARITY.
 FT MOD_RES 165 165 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT SPOEDENCE 841 AA; 94571 MW; ODD31920DDE7E58 CRC64;

Query Match 13.6%; Score 847.5; DB 1; Length 841;
 Best Local Similarity 28.1%; Pred. No. 5, 9e-25;
 Matches 263; Conservative 160; Mismatches 316; Indels 197; Gaps 34;

QY 4 YVLQKIGSGFGKALIVKSTEDGRQYIKETINSRMSKREESREVAVALNMKHPI 63
 DB 4 YVLQKIGSGFGKALIVKSTEDGRQYIKETINSRMSKREESREVAVALNMKHPI 63
 QY 6 YCYLRVVGKSGYGEYLVKHDRGQYIKETINSRMSKREESREVAVALNMKHPI 65
 DB 6 YCYLRVVGKSGYGEYLVKHDRGQYIKETINSRMSKREESREVAVALNMKHPI 65
 QY 64 VQYRESFE-ENGSLVYIMDYCEGDLFRKINAKGVLFQEDQIILDMFVQICLALKVH 122
 DB 64 VQYRESFE-ENGSLVYIMDYCEGDLFRKINAKGVLFQEDQIILDMFVQICLALKVH 122
 QY 123 KIILHROIKSONIFLTGDTGVQGLDFGIARVLSVTELARTCIGTPYILSPICENKPY 182
 DB 123 KIILHROIKSONIFLTGDTGVQGLDFGIARVLSVTELARTCIGTPYILSPICENKPY 182
 QY 183 KSDVWALGCVLYELCTLKAHAFAGSMKNLVKTIISGSPVSLHVSVDLSVQLFK 242
 DB 183 KSDVWALGCVLYELCTLKAHAFAGSMKNLVKTIISGSPVSLHVSVDLSVQLFK 242
 QY 242 KSDVWALGCVLYELCTLKAHAFAGSMKNLVKTIISGSPVSLHVSVDLSVQLFK 242
 DB 242 KSDVWALGCVLYELCTLKAHAFAGSMKNLVKTIISGSPVSLHVSVDLSVQLFK 242

```

OY 243 PRDRPSVNTLEKGFIAKRIEKLSPOLAEFECLTKFSKFSOPIPAKRPAASONSISV 302
DB 246 PEERVSATILROYIKROISLFFLE-----ATKIKTSKNNT-- 281
OY 303 MPACKITRPAKAKYCIPLATKKYGDKTLHEKKPL-----OKHKOHNPEKRVNNGE 353
DB 202 -----KNGDS-----QSKPRATVVSGEASNEHVHPOP----- 310
OY 354 ERKRISEPAARKRL-----EFIEKKKOKODII---SLMAEQMKRO-----EKEHLERIN 402
DB 311 -----LSSFGSOTYIMGSKCLSQEKPRASGLLTKSPASTLKHKKCOLNSTLTLATISSVN 366
OY 403 RAREGMRNVLSAGSGSVNAPFLGSGGTAPSSFSRQOYEYHNAIFDOMOQOAEADNE 462
DB 367 -----IDLPAKGRDSVSDGFOEN---QPRYLDASNEGLGICSI--SOVEEMLODNT 415
OY 463 AK-----WKREIYGRGLPEROKG--QLAVERAKOYEELQKREKRAOKKARAEQSHM 511
DB 416 KESAPENLIPMMSDIDYTGKKNPVPKPLQPLKEQPKDOSL-----ALSPKLECSGTI 470
OY 512 VYLARLROILQNFNERQOIKAKLGEKKEKANHSEGOSEADMRKRIE--SLKANAANA 570
DB 471 LAHSNL-----RLIGSSDSFASASRYAGITGVCHNAODVAGECILIEKGRHIDPLPHNSG 527
OY 571 RAAVLKEOLERKREKAYEREKRVNEHLVAKGVKSSDVSP--LGOHETGSGSPSKOOMS 628
DB 528 SEPSLSROROKRRROQHR--GEKROVRHDLFAFOESPPRLPSHPIVGK----- 576
OY 629 VISVTSALKEVG-----VDSILDTRETSEEMQTNNAISSKRELLRLNENIKAOEDK 683
DB 577 -VGVSTOKEANORRVYTVGSVSSR--SSEMSKDRPLSARE--RRRLKO--SOEEMS 629
OY 684 GMONLSDFEIYNVEDAKEHEK-----SVSSDKKWEAGOLVPLDELITD 732
DB 630 SSGPSVRKASLTVAGPGKQOEEDQPLPARLSSDQSVTOEKQIHC-----LSEDELS 684
OY 733 TSESTERTYGEVITKLGPNSSPRAMKSPFDVSLKILGAEIQLQLOTELL-----E 784
DB 685 TSTSDSDGDYGE-----GKQGTNEI-----NALVOLMTQTLKDKESCE 725
OY 785 NNTIRSEIS--PEGEKYKPLINGEKRVOCISHEIN---PSAIVDSVETKSEFESA-- 836
DB 726 DVPANVSEFEKLHRRYRDTLLHGRVABEAEHIFKELPSALM--FGSEKIRLIVYLR 783
OY 837 -----SPQMSLKEGNELEPPDDETEI--LOEPESG 864
DB 784 TDVIRGLGVOLLEOVYDILLEDEDEPREVRLKHNHG 819

RESULT 3
NEK3_MOUSE STANDARD: PRT: 511 AA.
AC 09K0AS; 0920X9; 40. Created
DB 16-OCT-2001 (Rel. 40, Last sequence update)
DB 16-OCT-2001 (Rel. 40, Last annotation update)
DB 16-OCT-2001 (Rel. 40, Last sequence update)
DB Setline/lineone-protein kinase NEK3 (EC 2.7.1.-) (NIMA-related
DE protein kinase 3).
GN NEK3.
OS Mus musculus (Mouse).
OC Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OC NCBI_taxid=10090;
RN 111
RP SEQUENCE FROM N.A.
RX MEDLINE=99240743; Pubmed=10224116;
RA Tanaka K., Nigg E.A.;
RT "Cloning and characterization of the murine Nek3 protein kinase, a
RT novel member of the NIMA family of putative cell cycle regulators.";
RT J. Biol. Chem. 274:13491-13497(1999).
RN 121
RP SEQUENCE FROM N.A.

```

```

RA Chen A., Yanai A., Arama E., Kilfin G., Motro B.;
RT "NIMA-related kinases: Isolation and characterization of murine nek3
RT and nek4 cDNAs.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
CC -1- FUNCTION: KINASE THAT MAY PLAY A ROLE IN MITOTIC REGULATION.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC NIMA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL: AP093416; AAD20986.1; -.
CC EMBL: AF099066; AAD16286.1; -.
CC DBSP: P24941; ICKP.
CC MDL: MGI:1344371; Nek3.
CC InterPro: IPR00719; Puk_kinase.
CC InterPro: IPR002290; Ser_thr_kinase.
CC InterPro: IPR001245; Tyr_kinase.
CC Pfam: PF00069; pkinase.1.
CC PRINTS: PR00109; TYRKINASE.
CC SMART: SM00220; S_TKC.1.
CC PROSITE: PS00107; PROTEIN KINASE_ATP; 1.
CC PROSITE: PS00108; PROTEIN KINASE_ST; 1.
CC PROSITE: PS50011; PROTEIN KINASE_DOM; 1.
CC TRANSFERASE: Serine/threonine-protein kinase; ATP-binding; Mitosis;
CC K nuclear protein; Phosphorylation; Cell cycle; Cell division.
CC DOMAIN 4 255
CC FT NE_BIND 10 18
CC FT BINDING 33 33
CC FT ACT_SITE 125 125
CC FT RES 163 163
CC FT CONFLICT 239 239
CC FT CONFLICT 342 343
CC SEQUENCE 511 AA; 57222 MW; DE6D6C053C7302F CRC64;

Query Match 13.3%; Score 833; DB 1; Length 511;
Best Local Similarity 37.4%; Pred. No. 1.2e-24;
Matches 203; Conservative 78; Mismatches 174; Indels 88; Gaps 16;

OY 1 MEKVRLQKIGESGFKALIVYKSTEDGROYIKETINISMSSKEREBSREAVLANMKH 60
DB 1 MDNTYVLRVIGQSGFALVLYLOESSNQTFAKET--RLKSDPTYSKRAVLANMKH 57
OY 61 PNIVQRESEFENGSLYIVNDYCEGGDLFRINAQGVLFQEDQILDFVYICLAKHVN 120
DB 58 PNIVAFKESFEABEGYLYIVMEYCDGDLMDKROKGNLFEDTILNFIQICLGVNHH 117
OY 121 DRKILHRDIKQSNIFLTKDGTVOLGPGIARVINSTVELARTCTPTPYLSPEICENKPY 180
DB 118 KRVYLHRDIKSNVFLHNRKVLGPGSRRLSSPMACVYCTPIYVPELWENLAPY 177
OY 181 NKRSDVMAIGCVILECTLKHAFBAGSMKNLYKTISSGFPVSLAHYXVDLSVLSOLF 240
DB 178 NKRSDVMSLGLIVELCALNHPDANSKMKLILQGGIHPILPALYCKIQLGIYKQMLK 237
OY 241 RNPDRPVSINLEKGFIAKRIEKLSPOLIAE-----ECITFKFSKFSOPIPAKRP 293
DB 238 RNPDRPVSATLTLGRGSLAPVPCPLPQILIREYGOILDEIKITSPNMKKKQOSNRVGR 297
OY 294 ASGONSISVPAOKI-----TKPA-----AKYGIPLAKKYGDKRLH--EK 332
DB 298 ALGEANSAKMODEERKRCSTITELESTIGTIPANALGRARANP-----ESGNROHGSHT 353
OY 333 KPLQKHKAHOT--PEKRVN-----TGEERRKIS-----EAAKRRRLFEIKE 374
DB 354 SPASPRKRWERHGPSSNVEALEKASILTSSFAEDDGGSVKYAEENARQWV----- 407

```

QY 375 KKKKDDIISLMKAEOMKROEKERLERINRAREOGWRNVLASGGSGEKAPLGGCTIAR 434
 DB 408 REPEELLALSMKADSLQAFQF-----YTIYRGAESEFLGPI--SEDRASD 452
 QY 435 SSSSSRQGY---EHYHAFEDMOOQRAEDNE-AKKREI-----YGRGLPERQKGLAV 484
 DB 453 SYDGDLDSDVMDPEREPRLDEEDTDDEEDNENPDWVSELKKHYGDPG----PGGLLG 508
 QY 485 ERA 487
 DB 509 ERA 511
 RESULT 4
 NEK3_HUMAN STANDARD: PRT; 459 AA.
 AC 1536;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Serine/threonine-protein kinase NEK3 (EC 2.7.1.1) (NIMA-related protein kinase 3) (HSPK 36) (Fragment).
 GN NEK3
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE-Placenta;
 RX MEDLINE=94368699; PubMed=7522034;
 RA Schultz S.J., Fry A.M., Suetterlin C., Ried T., Nigg E.A.;
 RT "Cell cycle-dependent expression of Nek3, a novel human protein kinase related to the NIMA mitotic regulator of Aspergillus nidulans";
 RT Cell Growth Differ. 5:625-635(1994).
 RL (2)
 RN Cell Growth Differ. 5:625-635(1994).
 RP SEQUENCE OF 29-142 FROM N.A.
 RX MEDLINE=94100173; PubMed=8274451;
 RA Schultz S.J., Nigg E.A.;
 RT "Identification of 21 novel human protein kinases, including 3 members of a family related to the cell cycle regulator nima of Aspergillus nidulans";
 RT Cell Growth Differ. 4:821-830(1993).
 RL -1- FUNCTION: KINASE THAT MAY PLAY A ROLE IN MITOTIC REGULATION.
 CC -1- DISEASE: MIGHT BE CANDIDATE FOR USHER SYNDROME, BECAUSE OF ITS CHROMOSOMAL LOCATION.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC NIMA SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC EMBL: 229067; CAA82310.1;
 DB EMBL: 229067; CAA82310.1;
 DR HSSP: P24941; ICRP.
 DR MIM: 604044; ICRP.
 DR InterPro: IPR000719; Euk-pkinase.
 DR InterPro: IPR002290; Ser_thr-kinase.
 DR Pfam: PF00069; Pkinase.1.
 DR PROSITE: PS00109; PROTEIN_KINASE_ATP_FALSE_NEG.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST.1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM.1.
 KM Transferase: Serine/threonine-protein kinase; ATP-binding; Mitosis;
 KW Nuclear protein; Phosphorylation; Cell cycle; Cell division.
 FT NON_TER 1
 FT DOMAIN 1
 FT ACT_SITE 80 210
 FT ACT_SITE 80 80 BY SIMILARITY.

FT MOD_RES 118 118 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CONFLICT 30 31 IV -> LY (IN REF. 2).
 FT CONFLICT 140 142 SLG -> PSV (IN REF. 2).
 SQ SEQUENCE 459 AA; 52300 MW; DB506EAC30EAB49 CRC64;
 Query Match 12.7%; Score 790.5; DB 1; Length 459;
 Best Local Similarity 42.9%; Pred. No. 4e-23;
 Matches 156; Conservative 58; Mismatches 87; Indels 63; Gaps 5;
 QY 49 REEAVLVANMKHPNIVORYRESFEENGSLYIVMDYCEGDLFRKINAKGVLFQEDQIILDM 108
 DB 1 KREAVLSAKMKHPNIVAFKESFEAGHLYIVMEYCGDGLMKIKOKGKLPEDMIILNW 60
 QY 109 FVQICLALKVHDKRIILDRKSONIFLTKDGTVALDGFARVNSTVELARTCIGTPY 168
 DB 61 FTQKGLGVNHLKRRVILDRDKSKNIFLTQNGKVKLGDGFSARLLSNPMAFCTVGTPTY 120
 QY 169 YLSEICENKRYNKKSDIMALGVLYELCTLKHAFAGSMKNLYIKITSGSPVPLHYS 228
 DB 121 YVPEIENLPIYNNKSDIMSLGCLYELCTLKHAFAGSMKNLYIKVQGCISPLSHYS 180
 QY 229 YDLRSYLSOLFKNRPDRPVSNSILEKFTAKRIEKLSPQILAEFCLTKTFKFGSOP 288
 DB 181 YELQFLVKQFKRNPDRPVSNSILEKFTAKRIEKLSPQILAEFCLTKTFKFGSOP 225
 QY 289 PAKRPASGNSISVMPAKITKPAKYGIPLAKYGGDKLHKPKLPLOKHQAQTPPKR 348
 DB 226 -----YGEVLEIR-----NSKNTPRKK 245
 QY 349 VNTGEERRKISEEAKRRRLFEFEKEKKOKDQIISLMKAEOMKROEKERLERINRAREOG 408
 DB 246 TNSRIRIALGNBSATVQ-----EEQDRKSGHTDELSINE--NLVESALNRVNR-EEKG 297
 QY 409 WRNV 412
 DB 298 NKSIV 301
 RESULT 5
 NRKA_TRYBB STANDARD: PRT; 431 AA.
 ID NRKA_TRYBB
 AC 008942;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative serine/threonine-protein kinase A (EC 2.7.1.37).
 GN NRKA.
 OS Trypanosoma brucei brucei.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_Taxid=5702;
 (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-BATRO 164 / ISOLATE ISTARI;
 RX MEDLINE=93295429; PubMed=8515773;
 RA Vale M.J., Jr., Parsons M.;
 RT "A trypanosoma brucei gene family encoding protein kinases with catalytic domains structurally related to Nekl and NIMA";
 RT Mol. Biochem. Parasitol. 59:111-122(1993).
 RL -1- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphorylated protein.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC EMBL: L03778; AB59252.1;
 DR HSSP: P24941; IAO1.

DB 244 LNEITITRLMLNLDKVRHPSVEELLENPLIA-----DLVADE-QRRNLERGRGQ--L 290
 QY 291 KRPAQSGNSISVMPAQKITKPAKYGIPLAYKKYGDKKLHEKRPLOKHOAHOTPEKRYN 350
 DB 291 GEPEKQSDSSPVLSELT-----KEIQLOERERALKAREERLE 328
 QY 351 TGPE-----RRKISEE-AARKRRRL---EFIEKEK-----QKDDIISLMAREQKR----- 392
 DB 329 QKQDELQVARELAEDEKLRAENLKNYSILKEREFISLANPELNLPPSVIKKKYHFGS 388
 QY 393 QERERLERINRAREQ 407
 DB 389 ESKENIMRSENSESO 403

RESULT 7

NEK2_MOUSE STANDARD: PRT: 443 AA.
 AC 035942; 035959;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Serine/threonine-protein kinase Nek2 (EC 2.7.1.-) (Nima-related protein kinase 2).
 GN NEK2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN NCBI_TaxID=10090;

RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
 RC STRAIN=SWISS WEBSTER; TISSUE=Testis;
 RA MEDLINE=9730684; PubMed=9187143;
 RA Rhee K., Wolgemuth D.J.;
 RT "The NIMA-related kinase 2, Nek2, is expressed in specific stages of the meiotic cell cycle and associates with meiotic chromosomes";
 RT Development 124:2167-2177(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98243037; PubMed=9583679;
 RX Arama E., Yanai A., Kilfin G., Motro B.;
 RT "Murine NIMA-related kinases are expressed in patterns suggesting distinct functions in gametogenesis and a role in the nervous system";
 RT Oncogene 16:1813-1823(1998).
 RN [3]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC TISSUE=Brain;
 RX MEDLINE=98096235; PubMed=9434622;
 RA Tanaka K., Parvlien M., Nigg E.A.;
 RT "The in vivo expression pattern of mouse Nek2, a NIMA-related kinase, indicates a role in both mitosis and meiosis";
 RT Exp. Cell Res. 237:264-274(1997).
 RL -1- FUNCTION: PROTEIN KINASE THAT IS INVOLVED IN MITOTIC REGULATION. MAY HAVE A ROLE AT THE G2-M TRANSITION. MAY ALSO PLAY A ROLE IN MEIOSIS.

CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: MOST ABUNDANTLY EXPRESSED IN TESTIS. LOW LEVELS FOUND IN MID-GESTATION EMBRYO, OVARY, PLACENTA, INTESTINE, THYMUS AND SKIN. WITHIN THE TESTIS, EXPRESSION RESTRICTED TO GERM CELLS WITH HIGHEST LEVELS DETECTED IN SPERMATOCTES AT PACHTENE AND DILOTENE STAGES. ALSO EXPRESSED IN MEIOTIC PACHTENE OOCYTES.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC NIMA SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: U95610; AAB67973.1; -
 DR EMBL: AF013166; AAC35393.1; -
 DR EMBL: AF007247; AAB70470.1; -
 DR HSSP: P24941; 1CRP.
 DR MGD: MG1:109359; Nek2.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase.1.
 DR SMART: SM00220; S-TRK; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR Transferase: Serine/threonine-protein kinase; ATP-binding; Mitosis;
 KW Nuclear protein; Phosphorylation; Cell cycle; Cell division; Meiosis;
 FT DOMAIN 8 271 PROTEIN KINASE.
 FT NP_BIND 14 22 ATP (BY SIMILARITY).
 FT BINDING 37 37 ATP (BY SIMILARITY).
 FT ACT_SITE 141 141 BY SIMILARITY.
 FT MOD_RES 179 179 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 SQ SEQUENCE 443 AA: 51307 MW: DE09565C378307E1 CRC64;

Query Match 8.7%; Score 544.5; DB 1; Length 443;
 Best local similarity 27.7%; Pred. No. 4.6e-14;
 Matches 152; Conservative 85; Mismatches 182; Indels 129; Gaps 14;

QY 1 MEKTVRLQKIGSGSGKALVYKSTEDGRQYIKRINISRSKSEKRESREVVANMKH 60
 DB 5 VEDIEVLSIGTSGKCKIRKSDGKILVWKELDGSMTEVEKQMLSEVNLRLKLN 64
 QY 61 PNIVQYRESF--EENGSLIYMDYCEGDLFRKINAQV---LRQEDQILDMPVQTL 114
 DB 65 PNIVRYDYDIIDRTFTTYLYWEYCEGGLASVIS--KGRKQROYLEEVEVLVMTQLTL 122
 QY 115 ALKHVDRK----ILNRDKSONIFLTKDGTVOAGDGIARVNLSTVELARTICGPY 169
 DB 123 ALKECHRRSDGCHTVLHRDLKPAVNFILDSKNHNVLDGFGFLARILNHDTSFAKTFVCTPY 182
 QY 170 LSPFICENRPYNNKSDIYALGCVLVELCTLKHAEAGSMKNLVLIIGSFPPVSILHSY 229
 DB 183 MSPEQMSCLSYNEKSPDYSWLAELCLALMPPTAARNOKELAGKIREGFRRIPIRYSD 242
 QY 230 DLRSLSOLFKNPNDRPVSNIIEKGFIAKRIKFLSPOLIAEECLKFSFGSQPIR 289
 DB 243 GLNDLITRLMLFLKDVHRSVEELLENPLIA-----DMVAEE----- 278
 QY 290 AKRPASGNSISVMPAQKITKPAKYGIPLAYKKYGDKKLHEKRPLOKHOAHOTPEKRY 349
 DB 279 -----QRRNLERGRSGEP-----SKLPDSSPVLSELT-----KLKESQL 313
 QY 350 NTGEERRKISEEARKRRRLFEIEKEKQKQDI--SLMRAEQMKROEKRELERINRARE 406
 DB 314 QDEQQLARERDILDEKEREICIRERLEDEKLAEESIMNYSILKEHR----- 362
 QY 407 OGRNVLASGSGEYVAPLFGSGGTIAPSSFSFGQEHYALFDQMOQOARADNAKKK 466
 DB 363 -----LLCLAGGEPLDLP-----SSAKKKVYHFG-----ESKENYA--- 394
 QY 467 RELYGGLEPERQQLAVERAKOVEFLORREKREMONAKAREGMYLARLRIORFN 526
 DB 395 -----RSNSESYLAK-----SKRDLKRRLLHAQDLRQALADLE 429
 QY 527 EROQITAK 534
 DB 430 KNYOLKSR 437

RESULT 8
 ID NIMA_EMEI STANDARD: PRT: 699 AA.
 AC P11837;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)

QY	SEQUENCE	699 AA:	78912 MW:	50FDC86E249ABC99	CRC64:
Query Match	8.5%;	Score 530.5;	DB 1:	Length 699;	
Best Local Similarity	27.8%;	Pred. No. 2,4e-13;			
Matches 168:	Conservative 92;	Mismatches 206;	Indels 139;	Gaps 18	
2	EKKVYLQKIGSGPGKAIIVKSTEDGROYVIREINISMSKSKRESREKREVALANKMP	61			
Db	9 DKVEYLEIKIGSGSGIIIRKVRKRSDFILCRREINIMSKREEDLMEFNLISLEHP	68			
QY	62 NTIVQY--RESFEENGSLYIVMDYCEGDL-----FKINAKGVLFQDDOILDMFYIC	113			
Db	69 NIIVAYHHHHKASQDLYLYMEYCGGGDLNWKIKRKTKRYA-----EDDFWRLISQLY	124			
QY	114 LALKVH-----DKTLHDDIKSONIFLTKDGTVOLGD	146			
Db	125 TALLYRHGTGDPAYGNSNLGPAKPSGLKGAQAMTILHRDLKPEMIFGSDNPFVKGD	184			
QY	147 EGIAVLNLYSTVELATCIGPYIYSPETQENPKYNNKSIDWALGCVYELCTLKAHEAG	206			
Db	185 FGLSKLMHSH--DPAITYGPIFYMSPELCAAKYTLRSDIIVAGCIVELQCEHPFNAR	243			
QY	207 SMKNLYLKTISSPPVSLHAYSDLSVLSQFLKNRPDRPSVNSILEKFIKRIEKL	266			
Db	244 THIQLOKIREKFAFLPDPFSSLEKLVNASCILRVNPHDRPDTLLINTFVI-----	295			
QY	267 SPQILAEFCCLTFKSPGSDIPAKPRPASQNSISVMPAKITKPAKYGIPIAYKKYGD	326			
Db	296 --RLMRREVLNLS-----RAARKREATYOKAKKQDEVAFAK-----	331			
QY	327 KLIHEKKPLQKHQKHOPEKRVNTGERRKIEEFAARKRRLEFTEKKOKQDIIISLMK	386			
Db	332 -----LEKEKQOIRS-----BLEMSIRREWEVARELAIEDROVONELDKRRE	375			
QY	387 AEQMKR--QEKERLEIRINAREOGMRVNLVAGSGGEVKAFLPGSGGTAPDSFGSGOTE	444			
Db	376 CEVDYRAQGVVEEQRNANATYRDA--SLRSGSHSQM-----SSNSGDSFSSSD--	425			
QY	445 HYHAIPOMOQOQRAEDNEAKMKREIYTGRLPERQKGLAVEAKQVEEFLQKREANQK	504			
Db	446 -----ISQLSELEPTTKARK-----LPKKE--SREPTTSKIVY-----DSPMDI	463			
QY	505 ARAEGHWYIATRLRQIRLONFERQOLKAKIRGEKKAEHNSGCGSGSEADMRKKLESL	564			
Db	464 QMAEESPISLTL-----SLSPKISATYS--GKNIPAEGERKKRPKEPT	506			
QY	565 KAHAN 569				
Db	507 LAYSD 511				
RESULT	9				
NRKB_TREUB	NRKB_TREUB	STANDARD:	PRF:	431 AA.	
ID	NRK3428:				
DC	01-PCB-1994 (Rel. 30, Created)				
DT	01-PCB-1996 (Rel. 33, Last sequence update)				
DE	16-OCT-2001 (Rel. 40, Last annotation update)				
GN	putative serine/threonine-protein kinase B (Ec 2.7.1.37).				
OS	Trypanosoma brucei brucei.				
OC	Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.				
NCBI_Taxid:	5702;				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ISOLATE TREU66;				
RA	MEDLINE=93295429; PubMed=8515773;				
RT	Gale M.J. Jr., Parsons M.;				
CC	"A trypanosoma brucei gene family encoding protein kinases with catalytic domains structurally related to nek1 and NIM1."				
CC	RT Mol. Biochem. Parasitol. 59:111-122(1993).				
CC	-1- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphoprotein.				

```

CC CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC CC at the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: L03777; AAB59253.1; -
CC CC HSSP: P24941; IAQ1.
CC CC InterPro: IPR000719; Euk_pkinase.
CC CC InterPro: IPR001849; PH.
CC CC Pfam: PF00169; Ser_Pkinase.
CC CC SMART: SM00220; S_TKc.1.
CC CC PROSITE: PS00107; PROTEIN_KINASE_ATP, 1.
CC CC PROSITE: PS00108; PROTEIN_KINASE_ST, 1.
CC CC PROSITE: PS50011; PROTEIN_KINASE_DOM, 1.
CC CC Transferrase; Serine/threonine-protein kinase; ATP-binding.
CC CC TRANSFERASE; Serine/threonine-protein kinase; ATP-binding.
CC CC DOMAIN 20 279 PROTEIN KINASE.
CC CC NP_BIND 26 34 ATP (BY SIMILARITY).
CC CC BINDING 49 49 ATP (BY SIMILARITY).
CC CC ACT_SITE 147 147 BY SIMILARITY.
CC CC FT 431 AA; 48172 MW; F684B3EF9E9FE74 CRC64;
CC CC
CC CC Query Match 8.5%; Score 528.5; DB 1; Length 431;
CC CC Best Local Similarity 38.3%; Pred. No. 18e-13;
CC CC Matches 103; Conservative 62; Mismatches 99; Indels 5; Gaps 2;
CC CC
CC CC 3 KYVRLQKIGESFGKALIVSTEDGRQYIKENIRMSKRESRREAVLANMKHPN 62
CC CC 19 KYLNKGIVIGSGYGEVVAERVDGSLCAKAWDLKMRKRKRYAQSIEIKYPTNCNHPN 78
CC CC 63 IVQYREFSENGSLYIVMDYCEGGDLFKRI---NMOKGVLRDQDILDMFOVCIATKRV 119
CC CC 79 IIRYIEDHEENDRLIYMEFADSGNLDEQIKPMGTGDARVPEHALLFLQDLCLALDYI 138
CC CC
CC CC 120 HDRKILHRDIKSONIFETKDGTVQGLDGIARVLSNIVE--IARCTGTPYLSPECEN 177
CC CC 139 HSHMLHRDIKSNANVLTFTGLVTKLDFGSHQYEDIVSGVASFCTPIYLAPELMNN 198
CC CC 178 KPYNNKSDIWAICGVLYELCTLKHAFEGAGMKNLVLTISGSPVSLHYSDLRSLYSQ 237
CC CC 199 LRYNKKADVYSLGLVLYELTGMKKPFSSANLKLGMSTKVLATYAPLPSFSEFKRYVDG 258
CC CC
CC CC 238 LFRNRPDRPSVNSILEKGFIAKRIEKL 266
CC CC 259 ILVADPNDRPSVNEFQIYINKGLFLV 287
CC CC
CC CC RESULT 10
CC CC GINA_YEAST STANDARD: PRT; 1142 AA.
CC CC
CC CC 01-NOV-1997 (Rel. 35, Created)
CC CC 01-NOV-1997 (Rel. 35, Last sequence update)
CC CC 30-MAY-2000 (Rel. 39, Last annotation update)
CC CC Serine/threonine-protein kinase GINA (EC 2.7.1.1.-).
CC CC GINA OR YDR507C OR D9719.13.
CC CC Saccharomyces cerevisiae (Baker's yeast).
CC CC Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
CC CC NCBI_taxonomy:4932;
CC CC
CC CC SEQUENCE FROM N.A.

```

```

RX RX MEDLINE:99030835; PubMed:9813093;
RA Longtine M.S., Fares H., Pringle J.R.;
RT "Role of the yeast Glnp protein kinase in septin assembly and the
RT relationship between septin assembly and septin function.";
RL J. Cell Biol. 143:719-736(1998).
CC CC (2)
CC CC SEQUENCE FROM N.A.
CC CC RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E., Benito A.,
CC CC Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
CC CC Hanuicke-Smith S., Hyman R., Komp C., Laskari D., Lew H., Lin D.,
CC CC Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
CC CC Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
CC CC Wiant A., Yelton M., Botstein D., Davis R.W.,
CC CC Submitted (Aug-1995) to the EMBL/Genbank/DBJ databases.
CC CC -1- FUNCTION: MAY PLAY A ROLE IN SEPTIN ASSEMBLY.
CC CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CC NIMI SUBFAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC CC at the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: U33140; AAB75513.1; -
CC CC HSSP: U33057; AAB64949.1; -
CC CC SGD: S0002915; GINA.
CC CC InterPro: IPR000719; Euk_pkinase.
CC CC InterPro: IPR002290; Ser_Pkinase.
CC CC Pfam: PF00069; pkinase.1.
CC CC SMART: SM00220; S_TKc.1.
CC CC PROSITE: PS00107; PROTEIN_KINASE_ATP, 1.
CC CC PROSITE: PS00108; PROTEIN_KINASE_ST, 1.
CC CC PROSITE: PS50011; PROTEIN_KINASE_DOM, 1.
CC CC Transferrase; Serine/threonine-protein kinase; ATP-binding.
CC CC TRANSFERASE; Serine/threonine-protein kinase; ATP-binding.
CC CC DOMAIN 19 289 PROTEIN KINASE.
CC CC NP_BIND 25 33 ATP (BY SIMILARITY).
CC CC BINDING 48 48 ATP (BY SIMILARITY).
CC CC ACT_SITE 156 156 BY SIMILARITY.
CC CC FT 1142 AA; 129857 MW; EC16F4BB49DD811 CRC64;
CC CC
CC CC Query Match 8.0%; Score 502.5; DB 1; Length 1142;
CC CC Best Local Similarity 19.8%; Pred. No. 4.4e-12;
CC CC Matches 252; Conservative 215; Mismatches 475; Indels 331; Gaps 42;
CC CC
CC CC 8 OKIGESFGKALIVKSTEDGRQYIKET---NISMSSKEREES-----RREY 52
CC CC 23 ETLGIGSTGQVQLARNGSTGOEAIVKSKAVNTNGSTIVSGTTPDALPYGIENI 82
CC CC 53 AVLANKMKHPNIVQYREFSENGSLYIVMDYCEGGDLFKRI---NMOKGVLRDQDILDMFOV 112
CC CC 83 IIMKLNHPNVLRLYDWEENTDLYLVLEVAEKELFNLL-VERGL-PEHAIKIFFQOI 140
CC CC
CC CC 113 CLAKVHDKRLIHRDIKSONIFETKDGTVQGLDGIARVLSNIVE--IARCTGTPYLSPECEN 177
CC CC 141 IIGVSYCHALGIVARDKLPENLLDHRKYNKIKADFGMA-ALTEBEKLIETSGSGSHYAP 199
CC CC 173 EICENKPYNN-KSDIWAICGVLYELCTLKHAF--EAGSMKNLVLTISGSP-PVSLHYS 228
CC CC 200 EIVSGIYGFASDVVSCGVIILFALLTGLRIPDEEDGNIRTLTLVQGEFEMSSDDEIS 259
CC CC 229 YDLKSLVSQLFKRNPDRPSVNSILEKGFIAKRIEKL-----RIEKLSQLIAEE 274
CC CC 260 READDLIRKLILYDIPRKRIRKTRDILKHPLOKYPISIRDSKIRGLPRRDYTLPLSSNS 319
CC CC 275 FCLKTESK-----FSGSPTPAKRPASGNSISVPAQKIRTPAK-----YGIPLAYKK 323
CC CC 320 SIDATYLLQNLVYLLMGRDEPGIR-----EKLREPGANAEKTLVYALLVRFKC 365

```



```

OY 271 IAEFECLTKESKSGQPIPAKRPASGONSISVMPAQKITKPAKXGCIPLAKYKGDCKLH 330
DB 286 -----
OY 331 EKKPLQKHAQHOTPERVNTGERRKISEEAARRKRLFEIEKEKKOKDOJISLMKAEOM 390
DB 295 KEKEVEEFSHTLKKEETLN-----KRIE-----LDSKLSALETER-----SSIRAE-- 337
OY 391 KROEKEKLERINFAEQGRNVLSAGSGEYKAPLGSGETTAPSSPSSRQGYEHYAIIF 450
DB 338 -----IDASLRRE-----EYKA-----RLEI 354
OY 451 DOMOQOAEADNEAKMKREI-----YKNG-----LPEROGOLAVERAKOYEELQK 497
DB 355 DRIVAOEISLQOKEEOEYQAVAEALORHGRGPMFNSHGQSGSPSTATLVSDY----- 410
OY 498 REMOKKABEGHAYLRLROIRLQNFNERQOITKANL-----RGEKKKANHSEGOEGS 551
DB 411 -NLSSVSGGDDPSTDDIDIDISTESTDSDITKRIPTPPHRAQYSSAPAEVGLT 468
OY 552 ----HEADMRKKIESLKAHANANAYLKEQLERR--KEAYEREKKWEEHLVAKGV 603
DB 469 PMDIEAASPPTTISLSL-SPRKALTKAPTNPRAIFGEETSTDKSNWEVPRETEMI 527
OY 604 KSDVS-----PRLOHETG-----GSPKQOMRSV 629
DB 528 DSGDESEAEALVPSPKRTKSKNPFSTVTRPSRLNSOONSVLPIHGLRKSOTLAR 587
OY 630 ISVTALKEGVND-----SLTDRETSEMOKTNNAISKREILRLNENLKAEDEKG 684
DB 588 SKTVGVSIGOHPLRSAPSLRDKRPSPTRLSRHSVTVGVGRSLANNINNSNG 647
OY 685 MONLSDPTEIVNHDAREHEKESVSDRKKMAGQVLPDELTLDTSTSTERTHTVG 744
DB 648 SDASSVTYSNITVTRGLKRMSTCD-----ESSFSQOQ----- 682
OY 745 EVIKLPGSPRRAMKSPDSDVLKLTIGAELOLOTELENTTINSEISPE-----GE 797
DB 693 -----NNOPOOSLPAPP-----LKKIG-----LMAAKNIGSSLVEHLHQAAGR 723
OY 798 KYKPLTNGEKVOCI-SHEINSAIVDSVETKSPFESAPOMLAKIEGNLEPPDD 853
DB 724 PISAIISNEAKLAFKEHATIASAVDS-----SSSSSSSSQSQSLPTPRPSOPETD 775

RESULT 12
STIO_MOUSE
ID STIO_MOUSE STANDARD: PRT: 966 AA.
DC 025098;
DB 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Serine/threonine-protein kinase 10 (EC 2.7.1.37) (Lymphocyte-oriented
GN SKI10 LOK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eulalia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=thymus;
RX MEDLINE=97426413; PubMed=9278426;
RA Kurumochi S; Moriuchi T.; Kuida K.; Endo J.; Semba K.; Nishida E.;
RA Karsayama H.;
RT FLOK is a novel mouse STE20-like protein kinase that is expressed
RT predominantly in lymphocytes;
RT J. Biol. Chem. 272:32678-32684(1997).
CC -1- FUNCTION: CAN ACT ON SUBSTRATES SUCH AS MYELIN BASIC PROTEIN AND
CC HISTONE IIA ON SERINE AND THREONINE RESIDUES.
CC -1- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphorylated
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN LYMPHOID ORGANS
CC SUCH AS SPLEEN, THYMUS, AND BONE MARROW.

```

```

CC -1- PTM: AUTOPHOSPHORYLATED.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC STE20 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@1sb-sib.ch).
CC -----
CC EMBL: D89728; BAA24073.1;
CC HSP: P24941; ICKP.
CC MGD: MGI:1099439; SKI10.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR002290; Ser_thr_pkinase.
CC Pfam: PF00069; pkinase; 1.
CC SMART: SM00220; S_TKc; 1.
CC PROSITE: PS00107; PROTEIN KINASE, ATP, 1.
CC PROSITE: PS00108; PROTEIN KINASE, ST, 1.
CC TRANSFAC: TF00011; PROTEIN KINASE, DOM, 1.
CC KW Phosphatase, Serine/threonine-protein kinase; ATP-binding;
CC phosphorylation; Coiled coil.
CC DOMAIN 1 234
CC FT DOMAIN 538 596
CC FT BIND 749 883
CC FT BIND 749 883
CC FT BIND 749 883
CC FT BIND 749 883
CC ACT SITE 157 157
CC ACT SITE 157 157
CC SEQUENCE 966 AA; 111992 MW; 7115EAC01032BF94 CRC64;

Query Match 7.5%. Score 468.5; DB 1; Length 966;
Best Local Similarity 21.8%; Pred. No. 6,6e-11;
Matches 179; Conservative 146; Mismatches 294; Indels 201; Gaps 28;

OY 9 KIGEGSFGKALYKSTEDGRQYIVKEINISMSSKEREESREAVLANMKHPNYORE 68
DB 41 ELGDGAFGVYKAKKKEGALAAKAYETK--SEELIEDYIEIETLACDHPHYTKLG 98
OY 69 SPEENGSLYIWDYCEGGDLFK-RINOKGVULOEOQILDMVOICLAKHVHDKRIIHR 127
DB 99 AYYDGKLMIMEPCPGAVDAIMLEDRGL--TEFOIYVCKOMLEALNLFGRKRIIHR 156
OY 128 DIKSONIFLTGDTVOVGDFGIARVNSTVELARCTIGPYIISPEI--CE--NKPNN 182
DB 157 DLKGNVLMTELEGDRIADFGVSANKLTKLQKRDSPICGTPYMAPEVULCEMKDAPDY 216
OY 183 KSDIWAIGCVYEECTLKHAPEAGSMKNLVKTIISGFPFVSL--HSYDLRSIVSOLF 239
DB 217 KADIVSLGILIEAQAQIEPHHELNPMRVLL-IAKSDPPTLLTPSKSVFERDPLKAL 275
OY 240 KRNPRDRSVNSILKEGFIARKIEKFLSPOLIAE----- 273
DB 276 DKNETPRSAOULLQHFYSRVSNKRLRELVAEKAEMEIEEDGRDGEEDAVDAVP 335
OY 274 -----EFCCLTKESKSGQPIPAKRP-----ASGONSI-SVMPAQK 307
DB 336 PLVNHOTDSANVTOPSLDSNKLLODSSTPLPSPQOEPVNGPCSPGDPLQTTSPADG 395
OY 308 ITRPAKYGIPLAAYKKYGDKKLHEKKPLQKHAQHOTPERKRVNNGEERRKISEEAARR 367
DB 396 LSKNDNDLAKVPLRK-----SRPLSMARIMDEKQIIPODEP--PSPAASKSQ 444
OY 368 LLETLEKKOKDOJISLMKAEOMKROEKERLERINFAEQGRNVLSAGSGEYKAPLFG 427
DB 445 -----RANO-SRPNSSALETL-----GGEA-----LT 465
OY 428 SGGTAPSSFSRSGOYEHYHALFDMOQOAEAD--NEAKMKREIYVGNGLP-----EROKG 480
DB 466 NGSLIEPSSVT-----PSHSKRASDCSNLSTESMDYGTSLADSLKNEKG 512

```

481 QLAVERAKVEEFLORR-----EAMONKARAEHIVYLAR--LROIRLON 524
513 SLISGSKLHNTKTRFRVVDVEVSITTSKIISEDEKDEERKLEIRLOELRLQ 572
525 FNE-----ROQIKAKLRGKKEKKAHNSGEGSEADM-----556
573 KEERNOTOLSKHELOLEQMHRKRFEOELNNAKKFYDVELEHKOQOYKMEODHSV 632
557 -RRKISLKAHANARAALVEQLERKREKAEYERKKWEHLAKGVKSSDVPPLGOH 615
633 RRKEKAKIRLEODRDYAKFOELKQKKEVSEVLEKPRQO-----RKESMKOMEH 686
616 ETGSGSPSKQM-RSVISVTSALKVEGVDSLDTRTSEMOKTNNATSKREILR- 670
687 -SOKKORLDRDFVAKQKEDLELAKRKLTTEHNR--EICKERCKLSKQELLDRDEA 740
671 ---RLNENLKAODEKGMQNLSTFEINVEDAKHEKER 707
741 ALMEHEHQLQERHQLVKQOLKQDFLQHRDLKHEKER 780
RESULT 13
ST10_HUMAN STANDARD: PRT: 968 AA.
AC 094804; G9UW4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine/threonine-protein kinase 10 (EC 2.7.1.37) (lymphocyte-oriented
kinase).
GN SK10 OR LOK.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=99216434; PubMed=10199912;
RA Karanouchi S., Matsuda Y., Okamoto M., Kitamura F., Yonekawa H.,
RA Karanouchi S., Matsuda Y., Okamoto M., Kitamura F., Yonekawa H.,
RT Molecular cloning of the human gene SK10 encoding lymphocyte-
oriented kinase and comparative chromosomal mapping of the human,
RT mouse, and rat homologues.*;
RU Immunogenetics 49:369-375(1999).
RU SEQUENCE OF 814-968 FROM N.A.
RE TISSUE-Tests;
RA Blocher H., Boecher M., Brandt P., Mewes H.-W., Gaassenhuber J.,
RA Klemann S.;
CC Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
CC HISTONE. ITA ON SERINE AND THREONINE RESIDUES (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + a protein - ADP + a phosphoprotein.
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN LYMPHOID ORGANS.
CC -1- PIM: AUTOPHOSPHORYLATED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC STE20 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB015718; BAA35073.1;
CC EMBL: AL133081; CAB61400.1;
CC HSSP: P24941; 1HCL.
CC MIM: 603919;
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR002290; Ser_thr_kinase.
CC InterPro: IPR001245; Tyr_pkinase.

DR Pfam: P00069; Tyrosine kinase.
DR PRINTS: P00109; TYRKINASE.
DR SMART: SM00220; S_TKc.1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST.1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW phosphorylation; Coiled coil.
FT DOMAIN 36 294 PROTEIN KINASE.
FT DOMAIN 573 947 COILED COIL (POTENTIAL).
FT DOMAIN 750 884 GLN-RICH.
FT NP_BIND 42 50 ATP (BY SIMILARITY).
FT BINDING 65 65 ATP (BY SIMILARITY).
FT ACT_SITE 157 157 BY SIMILARITY.
FT ACT_SITE 157 157 BY SIMILARITY.
SEQUENCE 968 AA; 112134 MW; 15E245193EC53D CRC64;
Query Match 7.38; Score 455; DB 1; Length 968;
Best Local Similarity 22.18; Pred. No. 2.1e-10; Indels 164; Gaps 26;
Matches 177; Conservative 144; Mismatches 317;
9 KIGESFGKALIVKSTEDGRQYVIREINISMSKREBSREKRVAVLANMKRPVYORE 68
41 ELGDGAFGKRVYAKKKEFGALAAKVIETK--SEELDEIVLEILATCDHPYIVKLLG 98
69 SPEENGSLIVYMDYCEGGDLFK-RINAKGVLFQEOIILDMFVQICLAKVHDKIIR 127
99 AAYHDKIMIMTEFGPGAVAIMLEIDRGL--TEPIQVCKOMLEALNFIHSRIIR 156
128 DIKSNIFLTKDGVOLDGFIARVLSYELANTCIGTFYXISPI--CE--NKPYNN 182
157 DLKAGVMTLESDIRLDFGVSANLKTLOKRSFICITFYWAPVWCVETMKDPYDY 216
183 KSDIALGCVLYELCTLKAHFAAGSMKLYIKIISGFPPVSV--HYSYDLRSLSQGF 239
217 KADIMSDIGTLEKMAOIEPHEHLPMAVILK-IKADDPETLLTPSKWSVEFRDIKAL 275
240 KNPDRPVSNSILEKGFIAKRIEKLSPOLIAE----- 273
276 DKNEPTEPSAOLLEPFPVSSITTSKALRELVAEKKEVMEIEDGRDEEDDAVDAAS 335
274 -EPLCKTSFGSQPIPAKRSAGONSISWPAQ--KIRPAKVGIPLAYKKGDKK 328
336 TLENHTONSSEVPSPINDKPLESPSPPLAPSGSQDSVNEPCS-----QSGDSS 387
329 LHEKKPLQKHKAHQTEPKRVNTGEERKISEARARRLEFIKERKKOKDQIISLAKAE 388
388 LQTTSP-----PVVAPGNENGLAVPVPLKRSR-----PVSMDARIQVQAEK 428
389 QMKQEKERLERINARE--OGWRNVLSAGSGEYKAPFLSGGTIAPSSFSRGOYEY 446
429 QVADGGDLSPANRSQKASQSPNSALLETGSEKL-----ANGSLPEPPAQAAG---- 479
447 HAIDQMOOQARAENEAQKREI--YGRGLP-----EROKGQLAVERAKVEEFLORR- 498
480 -----PSKRSDSCSLSCISEMDYGTNLSTDLINKEMGSLSIKDPKLYKTKLRTRK 532
499 -----EAMONKARAEHIVYLAR--LROIRLONNE-----527
533 FVYDVEVSITTSKIISEDEKDEENHFLRQELRLQKEERNOTOLSKNHELOLE 592
528 -----ROQIKAKLRGKKEKKAHNSGEGSEADM-----RRKISLKAHANARAAL 574
593 QMKHFEELNAKKFPDTELENIERQOKOYKMEODHAVERARIRLEODRDTYR 632
575 LKEOKERKRAVREKKWEHLVAGVSSDVPPLGOHETGSGSPSKQM-RSVISV 632
653 FQELKLMKKRVKNEVELEPQO-----RKESMKOMEH-----TOKKOLDRDFVAK 701
633 TSLKEVGVDSLDTRTSEMOKTNNATSKREILR-----RLNENLKAODEKGM 685
702 QKEDLELAKRKLTTEHNR--EICKERCKLMKQELLDRDEAALMEHEHQLQERHQLVK 759

QY 686 QNLSDFEINVEDAKEHEK 707
 Db 760 QOLKDOYFLOHRLKRHEK 781

RESULT 14
 KKRL YEAST STANDARD; PRT: 1518 AA.

AC 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Probable serine/threonine-protein kinase YKL101w (EC 2.7.1.-).
 GN YKL101w OR YKL453.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OX Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 RN 1111;taxid=4932;
 RP SEQUENCE FROM N.A.
 RC STRAIN=288C;
 RX MEDLINE=94078677; PubMed=8256524;
 RA Paller C., Valens M., Puzos V., Fukuhara H., Cheret G., Sor F.,
 RA Bolotin-Fukuhara M.,
 RT "DNA sequence analysis of a 17 kb fragment of yeast chromosome XI
 RT physically localizes the MBI gene and reveals eight new open reading
 RT frames, including a homologue of the KINI/KIN2 and SNF1 protein
 RT Yeast 9:1149-1155(1993).
 RL -1 SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -NIM1 SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC or send an email to license@ebi-sib.ch).
 CC
 DR EMBL: X71133; CAA50456.1; -
 DR EMBL: Z28101; CAA81941.1; -
 DR PIR: S37928; S37928.
 DR PIR: S39084; S39084.
 DR HSP: P13362; 1FGK.
 DR SGD: S0001584; YKL101w
 DR InterPro: IPR000719; Euk-kinase.
 DR InterPro: IPR002290; Ser-thr-kinase.
 DR SMART: SM00220; S_TKc_1.
 DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE: PS00108; PROTEIN KINASE; 1.
 DR PROSITE: PS50011; PROTEIN KINASE; 1.
 KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
 KW ATP-binding.
 FT DOMAIN 81 369 PROTEIN KINASE
 FT NF-BIND 87 95 ATP (BY SIMILARITY).
 FT BINDING 110 110 ATP (BY SIMILARITY).
 FT ACT_SITE 239 239 BY SIMILARITY.
 FT SITE 239 239 BY SIMILARITY.
 SEQUENCE 1518 AA; 169592 MW; 803f84f7531241DD CRC64;

Query Match 7.2%; Score 452; DR 1; Length 1518;
 Best local similarity 20.2%; Pred. No. 4,3e-10;
 Matches 260; Conservative 199; Mismatches 449; Indels 380; Gaps 50;

QY 10 IGGSGGKALVKESTEDGROYIKRI-----NISRRSK 43
 Db 87 LGGSSGGRVLAKNMETSOLAIRKIVPKKAFVCHSNNNGITVPSYSSMTSNVSPSIA 146

QY 44 EREES-----RREVAVALANKHPNIVQYRESFENSATYIVNDYCEGDLKRRIN 93
 Db 147 SRHSNHSQTNPNYGIEREIVIMKILSHTNVALFVEMNSSELYLVLEIVDGGSLDYL- 205

QY 94 AOKGVLFQEDQIDLDFVQICLAKHVDRKILHRDIKSONIFL-TRDGYOAGDRCIARV 152
 Db 206 VSKGKL-PEREAHIFKQIVGVSCHSFNICHRLDKPENLLDKNNRRIKIADROM-A 263

QY 153 INSTVELARTCIGTPYYSPEICENKPPYN-NKSDYMLGCVLELCTLHAFAEASMKL 211
 Db 264 IELPNKILKTSCGSPHYASPEIVMGRCGPDVMSCGIVFALLTGLHPNDONIKKL 323

QY 212 VIKIISGFPPVSLAHYSYDLKSLVSOQLFKNNPDPSVNSIIEKGTAR----- 260
 Db 324 LKVOGSKGYOMPS-NLSSEARDILSKILVIDPEKRIITGELLKHPILKIKYDLPVNVKL 382

QY 261 -RIEKLPSLIIEFCLTFFSKFGSOPIPARPPASGNSISVMPAOKITPAKAGI- 317
 Db 383 KMRDNNARCSNSDHLNLNVS-----PSIVYLHSGGEIDSLRSIQI-LMGVS 433

QY 318 -----PLAYKKYGDKKLHEKKPLQKHKAQHPERKRVNTGERRKISEAARKRLFEI 371
 Db 434 RELITAKLQKPMSEKLFYSLLO-YKORHSI-SLSSSENKKSATSESVNEPRIYA 490

QY 372 EKE-----KKQDQIISLAKADQMRKQERLERINRAREGQWNVLSAGSGEVKAPF 425
 Db 491 SKTANNTGLSENNDVKTLSLE-----HSEDTSTVN-----ONNAITGVNTEINAPV 539

QY 426 LGSGGTTAPSSFSRQYEHYNAIFDOMOORAEDNEAKWKREIYGRLOPERQGLAVE 485
 Db 540 LA-----OKSOPSIN 549

QY 486 RAKQVEEFLQKREAMONKARAGHMVYL-----ARLQIRLQNFENRQOIKAK 534
 Db 550 TLSQPE-----SDAQAEDA-VTLPAIPFNASSRIFRNSYTSISSRSRSLR 597

QY 535 LREKKEANHSQEQEESQADN-----RRKIESLKAHANAARAVALKEQLEKKE 585
 Db 598 LMSRSLSLSTREKLVHNDENRPLPOLKPSRYSIRRAIHAAPSRTSKSIHSLRRNIA 657

QY 586 AYREKKVVEHLYAKGVSSVSPPLCOHEGSGSPSKQOMRSYISVTS-A-LKEVGV- 641
 Db 658 ATVAARITQ-----NSAKRSLYSLSQSTSKSLNLDILVAF 694

QY 642 DSSLTDTRETEEMOKT-----NNAISSKRETLRLNLENLAQED 681
 Db 695 DDLPLPKRPSAENVKSEPHLSDSDFEILCDQLFGNALD--RLLEEEDNEKERDT 751

QY 682 EKGQMN-----LSDTPEINVEDAKEH-----EKEKVSVDKKWEAGQVLPI- 726
 Db 752 ORQRODRTKSSADLTITIGVSTNKNENPEPTYLEKQNMSTKPSNNGSLSPFIE 811

QY 727 DELTLDTSFSTERTTGEVYIKLQPNQSPRRANKSTYDYLKILGAELOLOTELENT 786
 Db 812 KENTLSSSYLERQK-----PKNALSDITTNMKNKQEGRIEKTQ-- 854

QY 787 TIRSEISPEGEKRYKPLINGEKKYOCISHEINPSAIVDPETKSPSEASPMQSLKLG 846
 Db 855 --RQLOKKNDPRPSL-----KPIQHQ-----ELRVSLPNDQKPSLSIDPRR 896

QY 847 NLEEDDLETEL-----QEPSC--TNKDESI-----PCTIVDVISSE--K 884
 Db 897 NISQPVNSKVESLLOGLKFKKEPASHWTHRGSLFSEHVEDKPKVASDVSISSSYVP 956

QY 885 EKEKQASDRTTQDENEVED-----GVSSYVDLSDIHIERPNDQSHSKCVD 934
 Db 957 TTVATSSRDPVLAESSYTIQPKMLSLPSSPLNMTFKNLSDIADGD----- 1006

QY 935 KSVQPEPFHVVASEHLNLYVPOVQVOCSPESFAFRSHS--HLPPKNKKNSLLGL 991
 Db 1007 -----KHLSS-VPOQNS--RSVAMSHPLRQNSAKISLITRSNINAMNLSYKR 1048

QY 992 STGLFDANNPKMLTCSLPLSLKFLTMDVPIVGVQNDNLEIDENIKPEPSQSE 1051
 Db 1049 NOG-----SPGSYLSNDLDGIDSM-TFAVEIPI-----NTFTAQ 1081

[illegible]

Wed May 15 14:14:05 2002

us-09-783-320-4.rsp

Page 14

Db 874 -----ENIEDGAEKGTESKIAASLSDDDUKE 900

Search completed: May 15, 2002, 08:02:27
Job time: 268 sec

[illegible]

seq.document.location.block:	1628 bp	mRNA	linear	HTC 04-OCT-200
LOCUS	BC015147			
DEFINITION	Homo sapiens, clone IMAGE:4063861,	mRNA.		
ACCESSION	BC015147			
VERSION	BC015147.1			
KEYWORDS	HTC.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 1628)			
AUTHORS	Straussberg,R.			
TITLE	Direct Submission			
IDENTIFI	Submitted (01-OCT-2001) National Institutes of Health, Mammalian			
	Genome Project (http://www.genome.gov/HTC04OCT2001). Cancer			

REMARK	COMMENT
<p> MGC Project URL: http://mgc.ncl.nih.gov Contact: MGC help desk Email: cgabbs-remail.nih.gov Tissue Procurement: ATCC GNA Library Preparation: CLONTECH Laboratories, Inc. GFP Consortium (ILMU) </p>	

```

ncup:77"
contact: amandansystemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kettelman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/BLINt at: http://image.llnl.gov
Series: IMAGE Plate: 31 Row: P Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis

This clone has the following problem: frame shifted.
Location/Qualifiers
1..1628
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4063881"
/tissue_type="Bone marrow, chronic myelogenous leukemia"
/clone_id="NH_MGC_54"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"

BASE COUNT      549 a      287 c      360 g      432 t
ORIGIN

alignment_scores:
Quality: 2038.00      Length: 400
Ratio: 5.108          Gaps: 0
Percent Similarity: 99.750      Percent Identity: 98.750

alignment_block:
US-09-783-320-4 x BC015147      ..

Align seg 1/1 to: BC015147 from: 1 to: 1628

1 MetGAlLsTtYrVAlArGlEnGlnLysIlleGlYluGlYsPerhEgLYLY 17
|||||
419 ATGCAAGACATGTGTGACTACAGCAAGATGGAGCAAGGCTCATTTGGAAA 468
|||||
17 sAtAlleLeuVallySerThrGluAspGlyArgGlnTyrVallelySG 34
|||||

```

469	AGCCATTCTTGTTAAATCTACAGAAAGATGGCAGACACTATGTTATCAAG	518
34	LUUleasnIleSerArgMetSerSerLysGUATAGGUGUUGUSeTArgAG	50
519	AAATTACATCTCACAAGATGTCGCAATTAAGAAAGAAATCACAAGACA	568
51	GUUValAlaValIleuAlaMetLysHisProGsnIleValIdmPyrTr	67
569	CAGATCTCAGATTGGCCAAACATGAGCATTCCAAAATTGTTCCAGATAG	618
67	GUUGSerPheGUUGUUGUAGGlySerLeuTyrIleValMetAspTyrCysG	84
619	AAATTCATTTGAAAGAAAGAGGCTCTCTCTACATATGATATGATTTACTGTG	668
84	LUUGlyGlyAspLeuPheLysArgIleAsnAlaGlnLysGlyValLeuPhe	100
669	AGGAGGGGAGATCTGTTTAAAGCCAAATAAATGCTCAGAAAGGCGTTTGGTTT	718
101	GUGUAspGlnIleLeuAspTrpPheValGlnIleCysLeuAlaLeuLys	117
719	CAGAGGATCAGATTTTGGACGCGTTTGTACAGATATGTTTGGCCCTCGAA	768
117	SHISValHisAspArgLysIleLeuHisArgAspIleLysSerGlnAsnI	134
769	ACATGTACATGATAGAAAAATCTTCATCGAGACATTAATTAATTCACAAACA	818
134	IePheLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAla	150
819	TATTTTAACTAAAGATGACACAGTACAACTTGGAGATTTTGGAAATGGCT	868
151	ArgValIleuAsnSerThrValGlnLeuAlaArgThrCysIleGlyThrPr	167
869	AGAGTCTTAATAGTACTGTACAGCTGGCTCGAACTTGATCATTAGAGACCCC	918
167	CTGTyrIleuSerProGlnIleCysGlnAspLysProTyrAspAsnLys	184
919	ATACACTGCTGCACGTGAAATCTGTGAAAACAAACCTTACAAATATAAATA	968
184	GAAPbIleTrpAlaLeuGlyCysValIleuTyrGlnLeuCysThrIleLys	200
969	GTGACATTTGGGCTCTGGGGGTGGTCTTATGAGCTGTGTACACTTAATA	1018
201	HisAlaPheGluAlaGlySerMetLysAsnLeuValLeuLysIleIleSe	217
1019	CATGCTTTTGAACTGGCAGATATGAAAAAACCCTGACTGAAATATATATC	1068
217	TrpLysSerPheProProValSerLeuHisTyrSerTyrAspLeuArgSerL	234
1069	TGGATCTTTTCCACCTGTGTCTTTGCATATTTCTTAAGATCTCCGACGTT	1118
234	euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValaAn	250
1119	TGGGTCTCAGATTATTTAAAAAATAATCCTACGGATACACCATCAGTCAAC	1168
251	SerIleLeuGluLysGlyPheIleAlaLysArgIleGluLysPheLeuSe	267
1169	TCCATATTTGAGAAAGGTTTTTATACCACCAAGCATTAAGAAAGTTTCTGC	1218
267	rProGlnIleuAlaGlnGlnPheCysLeuLysThrPheSerLysPheG	284
1219	TCTCTACGCTAATTGCAAGAAATTTTGTCTAAAAAATTTTTCGAAATTTTG	1268
284	LysSerGlnProIleProAlaLysArgProAlaSerGlnLysHisSerIle	300
1269	GATCCAGCGCTATACCGCTAAAAAGACCGAGTTCCAGGACAAAACCTGCAT	1318
301	SerValMetProAlaGlnLysIleThrLysProAlaAlaLysPyrLysI	317
1319	TCTGTATGCTGCTGCTCAGAAAAATTACAAGCCTCCCTTAATATATGAT	1368
317	eProLeuAlaTyrLysLysTyrGlyAspLysLysLeuHisGluLysLysP	334
1369	ACCTTTAGCATATAGAAATATGGAGATTAATAAAATTTACACGAAAGAAAGAAC	1418

```

334  rqlglnlshslsglnalahlthprrglulysarqvalsn 350
|||||
1419  cactgcaaaaacataaacagcccatcamaactcagaaagagntgamt 1468
|||||
351  thrhgylugluarqarqylsileserqgluualaalaaqylsarqar 367
|||||
1469  actggagaaagaaagacagaaaatattctgagaaacagcaagaaomcag 1518
|||||
367  glengluhpheileglulysglulysglulnysapcnnlletlesrl 384
|||||
1519  gctggaaatttattgaaaanaagaaaagaaacaaagggctcgatttattcgt 1568
|||||
384  eumetllyalagluglmetllysarqglnglulysgluarqleugluarq 400
|||||
1569  taattgaagcctgaaacaaatgaaagccagaaaaaahaaaaaahaaaaa 1618
|||||
seq_name: gb_htc:AK010196

```

seq_documentation_block: 1515 bp mRNA linear HTC 19-JAN-2002

LOCUS AK010196 Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310076f08.NIMA (never in mitosis gene a)-related expressed kinase 1, full insert sequence.

ACCESSION AK010196 AK010196.1 GI:12845466

VERSION HTCC: CAP trapper.

KEYWORDS Mus musculus (strain:G57Bl/6J) adult male tongue cDNA to mRNA, clone:11b:RIKEN full-length enriched mouse cDNA library

SOURCE clone:2310076f08.

ORGANISM Mus musculus

REFERENCE Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciuromorphi; Muridae; Murinae; Mus.

AUTHORS Carninci,P. and Hayashizaki,Y

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 95279253

PUBMED 10349636

REFERENCE 2 (sites)

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to Prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 3 (sites)

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuana,T., Taahito,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.

TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE 4 (sites)

AUTHORS The RIKEN genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 665-690 (2001)

MEDLINE 5 (bases 1 to 1515)

PUBMED 11076861

REFERENCE 5 (bases 1 to 1515)

AUTHORS Aichi,H., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Matsuda,T., Baldorelli,R., Bono,H., Brownstein,M., Bull,C., Carninci,P., Fukuda,S., Fukunishi,T., Futuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hirotsu,T., Horii,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koyu,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,

TITLE

Nunazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,K., Sakai,K., Sano,H., Saeki,D., Schiraldi,L., Shibata,K., Shidota,Y., Shinagawa,A., Shiraki,T., Sobabe,Y., Suzuki,H., Tagami,M., Tawara,A., Takahashi,F., Tanaka,T., Teijima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasuniishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki.Y.

JOURNAL

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsukumi-ku, Yokohama, Kanagawa 230-0645, Japan (E-mail:genome-res@gsr.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAAGAGAACAATCCACAGACTCTTTTGTTCCTTTTTTTTTTTTTTTTNNN-3'] .cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGGAGAGCATCAGATTAAATTAATATCCCCCCCCCC[3]'. cDNA was cleaved with XhoI and SacII. Cloning sites, 5' end: XhoI; 3' end: SacII.
Host: SOLR

FEATURES

source	Location/Qualifiers
1..1515	/organism="Mus musculus"
/strain="CS7BL/6J"	
/db_xref="MGI:MGI:1902619"	
/db_xref="taxon:10090"	
/clone="2310076F08"	
/sex="male"	
/tissue-type="tongue"	
/clone_lib="RIKEN full-length enriched mouse cdna library"	
/dev_stage="adult"	
1..1515	
/gene="Nek1"	
1..1515	
/note="NIMA (never in mitosis gene a)-related expressed kinase 1 data source:MGI, source key:MGI:97303, evidence:ISS"	

BASE COUNT 512 a 280 c 373 g 350 t

ORIGIN

alignment_scores:
Quality: 2020.00 Length: 498
Ratio: 4.430 Gaps: 3
Percent Simularity: 91.566 Percent Identity: 82.129

alignment_block:

US-09-783-520-4 x AKO10196 ..

Align seg 1/1 to: AKO10196 from: 1 to: 1515

```
153 LeuaasSerThrValGluleAlaArgThrcysllecllythrProtyrtY 169
      CTTAATAATGACGTGGTGCGCTGCATAGGCCCT.CCAATCA 78
30 CTTAATAATGACGTGGTGCGCTGCATAGGCCCT.CCAATCA 78
169 rleueserproglulliecygsluasnlusprotyrasnansulserrsp 186
      CTTCGTACCCTGAATCTGTGAACAACCCCTTAATACATRAAACGTCAC 128
79 CTTCGTACCCTGAATCTGTGAACAACCCCTTAATACATRAAACGTCAC 128
186 lertpalaleuglucyvalleutyrgluLeucysthlleuylshsalal 202
```

502 naslnyslaarqlaqluylhismetvaltyrleuulaarqlueauavg 519
1076 GAATTAAGCCCGACCCGACGACGACCTGGTTATTGGCAAGACTGAGCC 1125
519 InileatrgleuGlnAsnPhesnslarGlnGlnIlelyslalaysleu 535
1126 AAATTAAGACTACAAATTTAATGAGCGGCACAGATTAAAGCCAAACTT 1175
536 ATGGlygluLysGlnGlnAlaAsnHisserGlnGlnGlnGlnLysSerG 552
1176 CGTGGTGAGATTAAGAGAGCTGATGGTTACCAAGGACAAAGCAACGCA 1225
552 uGlnAlaAspMetArgArgLysLysLysIleGlnSerLeuLysAlaHisAla 569
1226 AGAGACTACGTAGCGCTCAAAAAGATGGAGTCACTTAAGCGCGCAACAA 1275
569 spAlaATGAlaAlaValleuLysGlnGlnGlnGlnAlaArgLysArgLys 585
1276 ATGCACGCGCTGCTACTTAAAGACACAGCTGACGAGAAAGAAAGAAAGA 1325
586 AlatyrglnArgGlnLysLysValITTPGlnGlnHisLeuValAlaLysG 602
1326 GCTTATGAAAGAAAGAAAGAAAGATGAGAAACATTGGTGGCAGAG.. 1373
602 yAlaLysSerSerAspValSerProProlenGlyGlnHisGlnGlnArg 619
1374 .GTAAAGAGCTCAAGATGTCTCTGCTTCGCTTGAACCTCTTGAAAGAGGTG 1422
619 LysSerProSerLysGlnGlnMetArgSerValIleSerValIleSerAla 635
1423 GTTCTCCATCAAAACACAGAGGTAGACGCTGTCTTTCGTGTGCTCAACT 1472
636 LeuLysGlnValGlyValAspSerSerLeuThrAspThrArg 649
1473 TTGAAGAAGAGTGGCGCTGGATGGAGATTTAAGCTTAAGTACCCAG 1514
seq_name: gb_hc:AK009484

seq_documentation_block: 1515 bp mRNA linear HMC 19-JAN-2002

LOCUS AK009484 Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310022M10.N1MA (never in mitosis gene a)-related expressed kinase 1, full insert sequence.

ACCESSION AK009484

VERSION AK009484.1 GI:12844311

KEYWORDS HMC; CAP; trapper.

SOURCE Mus musculus [Strain:C57BL/6J] adult male tongue cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library

ORGANISM Mus musculus

REFERENCE 1 (sites)

AUTHORS Carninci,P. and Hayashizaki,Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2 (sites)

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 3 (sites)

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,Y., Nishi,K., Kitunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Tajima,T., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujimoto,S., Inoue,K., Togawa,Y., Iwasa,M., Ohara,E., Watahiki,M.

TITLE	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, J., Okazaki, Y., Murimatsu, M., Inoue, Y., Kita, A., and Hayashi-Izaki, Y. Riken integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
PMID	20530913
PUBMED	11076861
REFERENCE	4 (Riken)
AUTHORS	FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5 (bases 1 to 1515)
AUTHORS	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Araki, A.,

TITLE Direct Submission
JOURNAL Submitted (10-JUN-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT Please visit our web site (<http://genome.gsc.riken.go.jp/>) for

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'] GAGCAGACGAGATCCAGACGCTCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapted of sequence [5'] GAGACAGATTCCTGGATTAAATTAATATATCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

source	1. .1515 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="MGD:MG1:1905247" /db_xref="taxon:10090" /clone="2310022M10" /sex="male" /tissue.type="tongue" /clone.lib="RIKEN_full-length enriched mouse cDNA library" /dev_stage="adult"
gene	1. .1515 /gene="Nek1"
misc_feature	1. .1515 /gene="Nek1" /note="NIMA (never in mitosis gene a)-related expressed kinase 1 data source:MGD, source key:MG1:97303, evidence:ISS"
BASE COUNT	512 a 281 c 374 g 348 t
ORIGIN	

alignment_scores: Quality: 1906.50 Length: 500
 Ratio: 4.190 Gaps: 5
 Percent Similarity: 91.000 Percent Identity: 79.600

alignment_block:
 US-09-783-320-4 x AK009484 ..

Align seg 1/1 to: AK009484 from: 1 to: 1515

```

151 ArgValLeuAsnSerThrValGluLeuAlaArgThrCysTleGlyThrPr 167
    |||||.....|
29  CGAGTCTTAATAGT...ATGTAGAGTGCGCGAATGCTATGACGCTTAA 75
    |||||.....|
167 cTyrTLeuSerProGluTleGysGluAsnLysProTyrAsnLysS 184
    |||||.....|
76  ATACTATTGTTCAC...TGAATCTGTAACAAACAGCTTATACATTAA 122
    |||||.....|
184 eRAspIleTrpAlaLeuGlyCysVal...LeuTyrGluLeuCysThrLeu 199
    |||||.....|
123 GTGACATTGTAGGCTTTGGCGCTGCTCTTTATGAACTTACGTAACCTT 172
    |||||.....|
200 LysHisAlaPheGluAlaGlySerMetLysAsnLeuValLeuLysTleI 216
    |||||.....|
173 AAATCATCTTGAAGCTGAAACATGAAACCTGTACTGAAAGATTAAT 222
    |||||.....|
216 eSerGlySerPheProProValSerLeuHisTyrSerTyrAspLeuArgS 233
    |||||.....|
223 CTCGGATCTCTTCTCCCACTGCTCCACATTAATCTATGATCTCCCA 272
    |||||.....|
273 eRLeuValSerGlnLeuPheLysArgAsnProArgAspArgProSerVal 249
    |||||.....|
273 GCTTGCTGTCTCACTTATTAAAGAAATCTTAGGGATGACCAATCAGT 322
    |||||.....|
250 AsnSerIleLeuGluLysGlyPheIleAlaLysArgIleGluLysPhe 266
    |||||.....|
323 AACCTCCATTTGGAGAAAGGTTTATAGCTAAACGAATGAAAGTTTCT 372
    |||||.....|
266 uSerProGlnLeuIleAlaGluLupheCysLeuLysThrPheSerLysP 283
    |||||.....|
373 CTCCTCCAGCTTATGTGACAGAAATTTGTCTAAAGAACCTTCAAGT 422
    |||||.....|
283 heGlySerGlnProIleProAlaLysArgProAlaSerGlyLinsSer 299
    |||||.....|
423 TTGGACACAGCTCTCCAGTAAAGACCAATCAGACAGCAAGGTTC 472
    |||||.....|
300 IleSerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyr 316
    |||||.....|
473 AGTTCTTTTGTCCCTGCTCAGAAATCAAAACCTGCTAAATACG 522
    |||||.....|
316 ylleProLeuAlaTyrLysLysTyrGlyAspLysLysIleuHisGluLys 333
    |||||.....|
523 AGGCCTTTAACATATAGAGATATGAGATTA...AAGTTCTTGAGAAA 571
    |||||.....|
333 yspProLeuGlnLysHisLysGlnAlaHisGlnThrProGluLysArgVal 349
    |||||.....|
572 AACCAACCCCAAAACATTAACAGGCGCCCTCAATTCCTCCGTAAGATATG 621
    |||||.....|
350 AsnThrGlyGluGluArgArgLysIleSerGlnGluAlaAlaArgLysAr 366
    |||||.....|
622 AATTCGAGAGAGAGAGAGATTAATGCTAGAGAACACCAAAA...AAAA 670
    |||||.....|
366 gATCLeuGluPheIleGluLysGluLysLysGlnLysAspGlnIleLies 383
    |||||.....|
671 GAGGTTCGAATTTATGAGAAAGAAAGAAACCAAAAGCTCAG...ATTA 717
    |||||.....|
383 eRLeuMetLysAlaGluGlnMetLysArgGlnGluLysGluArgLeuGlu 399
    |||||.....|
718 GGTTCCTGAGAGCTGAGACATGAAAGCCCAAGACAGACACGCTTGAG 767
    |||||.....|
400 ArgIleAsnArgAlaArgGluGlnGlyTrpArgAsnValLeuSerAlaG 416
    |||||.....|
768 AGGTAATATGGCCAGGAGAACAGATGAGAGAAATGTTTAAGCGCTCG 817
  
```

```

416 yGlySerGlyGluValLysAlaProPheLeuGlySerGlyGlyThrIleA 433
    |||||.....|
818 TGAAGCGGTGAGATGAAGGCTCTTTTGGCATTTGAGAGGGCGTCT 867
    |||||.....|
433 lAProSerSerPheSerSerArgGlyGlnTyrGlyHisTyrHisAlaIle 449
    |||||.....|
868 CTCATCACCCTGTTCTCTCCAGGCGCAGTATGAACTTTCATTCGTCAT 917
    |||||.....|
450 PheAspGlnMetGlnGlnGlnArgAlaGluAspAsnGlnLysTyrTrp 466
    |||||.....|
918 TTGACCAATATCAGCGGTAGACAGACAGATATATAGACAGATGAA 967
    |||||.....|
466 sArgGluIleTyrGlyArgGlyLeuProGluArgGlnLysGlnLeuA 483
    |||||.....|
968 GGGGGAATCTATGTGCTGATGGCTCCAGAAAGCAAAAGAGCACTTAG 1017
    |||||.....|
483 lValGluArgAlaLysGlnValGluGluPheLeuGlnArgLysArgGlu 499
    |||||.....|
1018 CTGAGAGAGAGCCCAACCAAGTGAAGATTCCTACACCTTAACGAGAA 1067
    |||||.....|
500 AlMetGlnAsnLysAlaArgAlaGluGlnHisMetValTyrLeuAlaAr 516
    |||||.....|
1068 GCTATGCAATTAAGCCCGAGCCGAGACACGCTGTTTATTTGGCAAG 1117
    |||||.....|
516 gLeuArgGlnIleArgLeuGlnAsnPheAsnGluArgGlnIleLysA 533
    |||||.....|
1118 ACTAGGCAATTAAGACTCAAAATTTTATGAGCGCCACAGATTAAG 1167
    |||||.....|
533 lAlLysLeuArgGlyGluLysGlyGluAlaAsnHisSerGluGlyGlnGlu 549
    |||||.....|
1168 CCNAATCTTGAGTGAATTAAGAAAGCTGATGATCAACAAAGGACAAAGA 1217
    |||||.....|
550 GlySerGluGluAlaAspMetArgArgLysLysIleGlnSerLeuLysAl 566
    |||||.....|
1218 GCACCTGAGAGACTGACATGAGCTCAAAAGATGAGACTCTTAAGGC 1267
    |||||.....|
566 hHisAlaAsnAlaArgAlaAlaValLeuLysGlnGlnLeuGluAlaGly 583
    |||||.....|
1268 GCNAACCAATACAGCTCTCTCTACTAAAGACACGCGAGCCAAAAA 1317
    |||||.....|
583 rGlyGluAlaTyrGluArgGluLysLysValTrpGluGlnHisLeuVal 599
    |||||.....|
1318 GAAAGGAAGCTTATGAAGAAAGAAAGAAAGTATGGAAGAAACATTGGTG 1367
    |||||.....|
600 AlAlaLysGlyValLysSerSerAspValSerProProLeuGlyGlnHisG 1414
    |||||.....|
1368 GCGAGG...GTAAAAAGCTCAGATGTTCTCTGCTTGTGAACCTCTTGA 1414
    |||||.....|
616 uThrGlyGlySerProSerLysGlnGlnMetArgSerValIleSerValT 633
    |||||.....|
1415 AACAGGTGTTCTCCATCAAGCAGCAGTGAAGCTGTCAATTCGTGA 1464
    |||||.....|
633 hSerAlaLeuLysGluValGlyAlaAspSerSerIleuThrAspThrArg 649
    |||||.....|
1465 CTTCACCTTGAAGAAATGAGGCGCTGATGGAAGTTTAACGTATCCAG 1514
    |||||.....|

```

seq_name: gb_est1:A1936517

seq_documentation_block:

LOCUS A1936517 692 bp mRNA linear EST 17-DEC-1999

DEFINITION WQ28hb08.x1 Soares,NFL.T GBC.SI Homo sapiens cDNA clone IMAGE:2329503 3' similar to SW.NEKL MOUSE P51954

ACCESSION IMAGE:2329503 3' similar to SW.NEKL MOUSE P51954

VERSION IMAGE:2329503 3' similar to SW.NEKL MOUSE P51954

KEYWORDS SERINE/THREONINE-PROTEIN KINASE NEKL ; mRNA sequence.

ORGANISM Homo sapiens

SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 692)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1721 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 444.

FEATURES

source
1. 692
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2329503"
/clone_lib="Scars_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker. Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung, NBHL19W, testis NM1, and B-cell
NCL-CCAP_GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Scars and M. Fatima Bonaldo.

BASE COUNT 187 a 138 c 128 g 238 t 1 others
ORIGIN

alignment_scores:
Quality: 1177.00 Length: 230
Ratio: 5.140 Gaps: 0
Percent Similarity: 99.565 Percent Identity: 98.696

alignment_block:
US-09-783-320-4 x AI936517/rev ..

Align seg 1/1 to reverse of: AI936517 from: 1 to: 692

```

142 VALGILNLEUGLYASPPHEGLYLEALAAAGVALLLEUNANSERTHYALGI 158
|||||
692 GTACACCTTGAGATTGGAAATGCTAGAGTCTTAACTAGTACTGTAGA 643
158 ULEUALAAGTHRCYSILEGLYTHRPROLYTYTYLLEUSERPROGLUILEC 175
|||||
642 GCTGCTCGAACCTTCATAGGAGCCCATCTACTCTGCACCTGAAATCT 593
175 YSGIUAASLYSPROTYRASNASNLYSSEASPILETSPALALEUGLYCYS 191
|||||
592 GTGAACCAACCTTACAAATATAAAGTGACATTGGCTCGGGGTGT 543
192 VALLEUTYRGULEUCYTHLEULYSHIALAPHEGLUALAGLYSERME 208
|||||
542 GTCCCTTATGACCTGTACACTTAACATGCTTTTAACTGACGACATAT 493
208 LYSASNLEUVALLEULYSILELIESERGLYSERPHROPROVALSERL 225
|||||
492 GAAAAACCTGGTACTGAAGATATAATCTGCACTTTTCCACCTGTGCTT 443
442 TGCATTATATCTATGATCTCCGAGTTGGTGCTCAGTTATTTAAAGA 393
225 EUHISTYRSETYRASPHEUARGSERLEUVALSERGINLEUPHELYSARG 241
|||||
242 AANPROAGASPARGPROSERVALASERILLEUGLIUYSGLYPHELI 258
|||||
392 AATCTTAGAGATAGACATGACATGCACTCATATTGAGAAAGTTTAT 343
258 EALIALYARGILEGLIUYSPHELEUSERPROGINLEULIAGLUGLUP 275
|||||

```

```

342 AGCCAAACGCAATGAAAAGTTCTCTCTCCACGCTATTCACAGAAAT 293
275 hecysleuylsrphselelyspheglyserginproleproalaly 291
|||||
292 TTGCTTAAAAACATTTTGGAAATTTGGATTCACACCTATACCGCTAAA 243
292 ATCPROLASERGLYGLNINSEVALMESERVALMETPROALAGLYSII 308
242 AGACACCTTCAGACCAAACTCGAATTTCTGTATGCTGCTCAGAAAAT 193
308 eThrlysprolalalytyrglyleproleualatyrlystysrg 325
192 TACAAAGCCGCGCTTAAATATGGAATACCTTACCATATTAAGAAATG 143
325 lyasphlyslleuhalisglulysproleuglnlyshisylglnala 341
142 GAGATTAATAATTTACAGAAAAGAAACCACTCCAAAACATTAACAGGCC 93
342 HISGLNTHRPROGLIYASARYVALANTRHGLYGLUGIUAARGLYSII 358
92 CATCAAACTCCAGAGAAGAGTAACTGAGAGAAGAAAGAGGAAAT 43
358 eSerGluAlaAlaArgLysArgLysLeuGluPhe 371
42 ATCTGAGAGAGCAGCAAGAAAGANGACTGGAATTTAT 3
seq_name: gb_est1:AI816275

```

seq_documentation_block:
LOCUS AI816275 816 bp mRNA linear EST 09-JUL-1999
DEFINITION au46d12.x1 Schneider fetal brain 00004 Homo sapiens cDNA clone
IMAGE:2517815 3', similar to SW:NEK1 MOUSE P51954
SERINE/THREONINE-PROTEIN KINASE NEK1, mRNA sequence.
ACCESSION AI816275.1 GI:5431821
VERSION EST
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 816)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Getseil, G., Jost, S.,
Kizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin,
J., Moore, B., Schellenberg, R., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
Other ESTs: au46d12.y1
Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: similarity on wrong strand
Seq primer: -400P from Gibco
High quality sequence stop: 424.

FEATURES

source
1. 816
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2517815"
/clone_lib="Schneider fetal brain 00004"
/sex="male"
/tissue-type="frontal lobe"
/dev_stage="5 months post-conception"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript SK (Stratagene);
Site_1: SbfI; Site_2: XhoI; Double-stranded cDNA was
prepared from human fetal brain tissue. 5' and 3' adaptors
were used in cloning as follows: 5' adaptor

alignment_block:
US-09-783-320-4 x BG706222 ..

Align seg 1/1 to: BG706222 from: 1 to: 719

```

850 GluProAspSerLeuGluThrGluLeuGluProSerGlyThrAs 866
|||||
7 GAACCTGATGATTGGAAAGAAATCTACAGAGCAAGCTGAAACAA 56
|||||
866 nlyAspSerLeuProGlyThrIleThrAspValTrrPleSerGlu 883
|||||
57 CAAAGATGAGACTTGGCTGCTACTATGCTGATGATGATGAG 106
|||||
883 lylGluThrLysGluThrGlnSerAlaAspArgIleThrIleGlu 899
|||||
107 AAAAAAGAAAGAAAGAAAGCTAGTCGAGATGAGATCACCATTGAG 156
|||||
900 AsnGluValSerGluAspGlyValSerSerThrValAspGlnLeuSer 916
|||||
157 AATGAGATTCTCGAAGATGAGAGTCTCGAGTACTGAGCAACCTAGTGA 206
|||||
916 pIleHsIleGluProGlyThrAsnAspSerGlnHsSerLysCysAsp 933
|||||
207 CATCATATGAGCGCTGAGAACCATGATCTCAGCACTTAATGTGATG 256
|||||
933 alAspLysSerValGlnProGluProPheHsIleValIleHsSer 949
|||||
257 TAGATTAATCTGTCGACCGGACCAACCATTTTCCATAGAGTTCATTC 306
|||||
950 GluHsIleAsnLeuValProGlnValGlnSerValGlnGlySerProG 966
|||||
307 GAACTTGAATCTGATGCTCAAGTCAATCAAGTCAAGTCAAGTCAAG 356
|||||
966 uGluSerPheAlaPheArgSerHsSerHsIleuProGlyAsnLys 983
|||||
357 AGAATCTTTCGATTCGATCTCAGCTGCACTTAAACCAAAAAATTA 406
|||||
983 snLysAsnSerLeuLeuIleGlyLeuSerThrGlyLeuPheAspAla 999
|||||
407 ACMAAATTCCTTCTGCTGATTCGACTTCAACGCTGCTGATGATGCA 456
|||||
457 AACCCAAAGATGTTAAGACATGTTCACTCCAGATCTCAAAAGCTGT 506
|||||
1000 AsnProLysMetLeuArgThrCysSerLeuProAspLeuSerLysLeu 1016
|||||
1016 eArgThrLeuMet..AspValProThrValGlyAspValArgGlnAsp 1032
|||||
507 CAGAACCTTATGGCATGTTCCACCGTAGAGATGTTGCTCAAGACAT 556
|||||
1033 LeuGlu..IleAspGluIleLysAsp...GluAsnIleLysGluGly 1048
|||||
557 CTTCGACATATGATGACTTCGACAGATGAACCTTAACAGAGACCTT 606
|||||
1048 eArgSerGluAspIleValPheGluGluThrAspThrAspLeuGlu 1064
|||||
607 CTGATTCGAAAGACAT..GTGCTGAGAAACTGACACAGATTTAACA 655
|||||
1065 LeuGlnAlaSerMetGluGlnLeuLeuArgGluGlnProGlyGlu 1081
|||||
656 CTGCGAGGCTCGATGGAACATCTTAGGAGAACCACTGTTGAAACA 705
|||||
1081 yfSerGluGlu 1085
|||||
706 ACAGTGAGAGAGAA 719

```

seq_name: gb_est2:BE881153

seq_documentation_block:

LOCUS BE881153 847 bp mRNA linear EST 20-OCT-2000

DEFINITION 601492024F1 NIH_MGC_69 Homo sapiens CDNA clone IMAGE:3894389 5'

ACCESSION BE881153

..

VERSION BE881153.1 GI:10329929

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

location/Qualifiers

1..847

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3894389"

/clone.lib="NIH_MGC_69"

/tissue.type="Large cell carcinoma, undifferentiated"

/lab.host="DH10B (phage-resistant)"

/note="Organ: Lung; Vector: pCMV-Sport6; Site:1: NCI; Site:2: SAT; Cloned unidirectionally. Primer: oligo dT. Average insert size 1.1 kb. Library constructed by Life Technologies"

BASE COUNT 313 a 129 c 225 g 180 t

ORIGIN

alignment_scores:

Quality: 1075.00 Length: 298

Ratio: 4.151 Gaps: 11

Percent Similarity: 86.913 Percent Identity: 81.879

alignment_block:

US-09-783-320-4 x BE881153 ..

Align seg 1/1 to: BE881153 from: 1 to: 847

475 ProGluArgGlnLysGluGlnLeuAlaValGluArgAlaLysGlnVal 491

|||||

3 CCAGAAAGGCAAAAGGCGACGCTGAGAAAGAGCTAAACAAAGTAGA 52

|||||

491 uGluPheLeuGlnArgLysArgGluAlaMetGlnAsnLysAlaArgAla 508

|||||

53 AGAGTCTCTCAGCGAAAGGAGAGCTATGCAAGATTAAGCTGACGCG 102

|||||

508 lUcGlyHsMetValTyrLeuAlaArgLeuArgGlnIleArgLeuGln 524

|||||

103 AAGACATATGCTTATCTGCGAAGAGCTGAGCAATAGACTAGAGAT 152

|||||

525 PheAsnGluArgGlnGlnIleLysAlaLysLeuArgGlyLysGly 541

|||||

153 TTCATATGAGCGCCACAGATTAAGCCAAACTTCGTGTAAGAAAGAA 202

|||||

541 uAlaAsnHsSerGluGlyGlnGluLysSerGluGluLysPheLys 558

|||||

203 AGCTAATCATTCGAGAGGACAGAGAGAGAGAGAGAGAGAGAGAG 252

|||||

558 rGlyLysIleGluSerLeuLysAlaHsIleAlaAspAlaArgAla 574

|||||

253 GCAAAAGAAATCGAATCACTGAGAGAGAGAGAGAGAGAGAGAGAG 302

|||||

575 LeuLysGluGlnLeuGluArgLysArgLysGluAlaLysGluArg 591

|||||

```

303 CTAAGAAAGCACTAGAAAGCAAGAAAGAGGCTTATGAGAGAGAAA 352
591 yslsvalatpuglulhislenuvalalalsgllyalyserserasp 607
353 AAAAACTGGAGAGAGATTTGGTGTAAAGACTTAAGAGTTCTGAT 402
608 valserpropoleuglylnhislenuvalalalsgllyalyserserasp 624
403 GTTTCCTCCAGCTTGGACACATGAAACAGTGGCTCTCCATCAAGCA 452
624 nglmetargservalleserlthrservalalalsgllyalyserserasp 641
453 ACAGATGAGATCTGTATTTCTGACTCAGCTTGAAGAGTTGGCG 502
641 alaspserleuthrsphrarglurhserglumetglulys 657
503 TGGACGATGATTAACTGATACCGGAACTTCGAAAGATGCCAAAG 552
658 ThrAsnAsn,AlaIleSerSerLysArgGluIleLeuArgLeuAsn. 673
553 AACAAACATGGCTATTTCAGAACGAGAAATCTTCGATTAAAT 602
674 ..GluAsnLeuLysAlaGlnGluAspGluLysGlyMet.GluAsnLeu 689
603 GAAATCTTTAAAGTCAAGAAAGATGAAAGAGACAGCAGATCTCTC 652
689 TASPThPheGluIle,AsnValHisGluAsp,AlaLysGlu,HisGlu 705
653 TGATACCTTTGAGTTAAGTTCATGAAAGATGCCAAAGAGCAGTGA 702
705 ysgLysSerValSerSerAspArgLysArgLysArgLysGln 721
703 AAGAAAAATTCAGTTCATCTGATGCAAGAGAGGAGGAGG..... 746
722 LeuValIleProLeuAspGluLeuThr..... 730
747 .....AGTTCCACTGCTGATTTCTGCTGAGGTTA 778
731 LeuAspThrSerPheSerThrArgLysHisThrValGluValI 747
779 ACTGATTAATCTTTCTTA.....AGGAGACTCAGAGGGGCA.... 815
747 lslsleuglyProAsnGlySerProArgArg 757
816 ..AGCTTAAAGGCTCAGGGGCTTCCAAACCG 845

seq_name: gb_est1:AV721713

seq_documentation_block:
LOCUS AV721713 641 bp mRNA linear EST 16-OCT-2000
DEFINITION AV721713 HTB Homo sapiens cDNA clone HTBAU07 5', mRNA sequence.
ACCESSION AV721713
VERSION AV721713.1 GI:10823477
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 641)
Gu, Y., Peng, Y., Song, H., Huang, Q., Yang, Y., Gao, G., Xiao, H., Xu, X.,
Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu
S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R.,
Chen, J., Chen, Z. and Han, Z.
Homo sapiens cDNA HTB clones
Unpublished (2000)
CONTACT: Zengqiang Han
Chinese National Human Genome Center at Shanghai
201203 P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzq@chgc.sh.cn
This clone is available at CHGC in Shanghai.

```

FEATURES

source

```

1..641
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HTB"
/clone_id="HTBAU07"
/issue_type="Hypothalamus"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site.1: EcoRI; Site.2:
XhoI"

```

```

BASE COUNT      209 a      126 c      134 g      169 t      3 others
ORIGIN

```

```

alignment_scores:
  Quality: 1057.00      Length: 208
  Ratio: 5.156          Gaps: 0
  Percent Similarity: 98.558      Percent Identity: 98.077

```

```

alignment_block:
US-09-783-320-4 x AV721713

```

```

Align seq 1/1 to: AV721713 from: 1 to: 641

```

```

829 LysSerProGluPheSerGluAlaSerProGluMetSerLeuLysleuG 845
3 GAAAGTCCGAGTTCAGTACGATCTCCACAGATGCTCATTTGAACGTGA 52
845 uGlyAsnLeuGluLysProAspAspLeuGluThrGluIleGlnGluL 862
53 AGGAAATTTAGAAAGACCTGATGATTTGGAAACAGAAATTTTCAAAG 102
862 roSerGlyThrAsnLysAspGluSerLeuProCysThrIleThrAspVal 878
103 CAAGTGGAAACAAACAAAGATGAGAGCTTCATGACACTTACTGATGTG 152
879 TrpIleSerGluGluLysGluThrLysGluThrGlnSerAlaAspArg 895
153 TGGATTACTGAGGAAAGAAAGAAACAAAGAACTGACGTGCGCATAGAT 202
203 CACCATTCAGGAAATGAGATTTCTGAAAGTGGAGTCTGACTACTGTGG 252
895 eThrIleGlnGluAsnGluValSerGluAspGlyAlaSerSerThrVal 912
912 spGlnLeuSerAspIleHisIleGluProGlyThrAsnAspSerGlnHis 928
253 ACCAACTTACGACATTCATATATAGCCCTGCAACCAATGATTTCCAGC 302
929 SerLysCysAspValAspLysSerValGlnProGluProPhePheHis 945
303 TCTAAATGTGATGTAGATTAATCTGTGCAACCAACCAATTTTCCATPA 352
945 sValValHisSerGluHisleuAsnleuValProGlnValGlnSerVal 962
353 GGTGGTTCATCTCGAACCTTGAACTTAGTCCCAAGTTCATTCAGTTC 402
962 IncYSerProGluLysSerPheAlaPheArgSerHisSerHisleuPro 978
403 AGGTTCACCAAGAAAGATCTTGATTTTCATTTTCATCTCAGCTTACCA 452
979 ProLysAsnLysAsnLysAsnSerLeuLeuIleLysLeuSerThrGly 995
453 CCAAAAAATTAAGAAAGAAATTCCTTGCTGATGGATGCTTCACTGCTC 502
995 uPheAspAlaAsnAsnProLysMetLeuArgThrCysSerleuProAsp 1012
503 GTTGTATGACNACCAACCAAGATGTAAAGCAAGATTCATCTCCAAATC 552
1012 euseLysLeuPheArgThrLeuMetAspValProThrValGlyAspVal 1028
553 TCTCAAGCTGTTCAGAACCTTATGATGATGTCACCGTANAGATGTT 602

```

1029 ArgGlnAspAsnLeuGluLeuLeuasp 1036
 |||||
 603 CGTCAAGACAACTCTGAAATAGAT 626
 seq_name: gb_est1:A1378221
 seq_documentation_block:
 LOCUS A1378221 642 bp mRNA linear EST 18-MAR-1999
 DEFINITION t655c01.x1 Soares_NbHMPU.S1 Homo sapiens cDNA clone IMAGE:2069472
 3' similar to SW:NEKL_MOUSE P51954 SERINE/THREONINE-PROTEIN KINASE
 NEK1; mRNA sequence.
 ACCESSION A1378221
 VERSION A1378221.1 GI:4188074
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 1 (bases 1 to 642)
 NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 691 Std Error: 0.00
 Seq primer: -400P from Glbco
 High quality sequence stop: 437.
 FEATURES
 source
 1. 642
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2069472"
 /clone_id="Soares_NbHMPU_S1"
 /tissue_type="Pooled human melanocyte, fetal heart, and
 pregnant uterus"
 /lab_host="DH10B"
 /note="Organ: mixed (see below); Vector: pTZ19-3D-Pac
 (Pharmacia) with a modified polylinker; Site 1: Not I;
 Site 2: Eco RI; Equal amounts of plasmid DNA from three
 normalized libraries (melanocyte 2NDHM, pregnant uterus
 NbHPU, and fetal heart NbH19M) were mixed, and ss circles
 were made in vitro. Following HAP purification, this DNA
 was used as tracer in a subtractive hybridization
 reaction. The driver was PCR amplified cDNAs from pools of
 5,000 clones made from the same 3 libraries. The pools
 consisted of I.M.A.G.E. clones 260232-265223,
 340488-343479, and 484488-489479."
 BASE COUNT 168 a 129 c 126 g 219 t
 ORIGIN
 alignment_scores:
 Quality: 1023.00 Length: 215
 Ratio: 4.895 Gaps: 2
 Percent Similarity: 97.209 Percent Identity: 94.419
 alignment_block:
 US-09-783-320-4 x A1378221/rev ..
 Align seg 1/1 to reverse of: A1378221 from: 1 to: 642
 159 LeuAlaArgThrCysIleGlyThrProTyrTyrLeuSerProGluIleCys 175
 |||||
 642 CTGGCTTGAACTTGATAGGACCCAAATCTCTGTCACCGAATCTG 593
 |||||
 175 sGluAsnLysProTyrAsnAsnLysSerAspIleTrrpAlaLeuGlyCysV 192
 |||||
 592 TAAACAAACCCCTTACATATATAAAGTACATTGGGCTCTGGGGTGTG 543
 |||||
 192 alLeuTyrGluLeuGlyThrLeuLysHis .AlaPheGluAlaGlySerM 208

542 TCCCTTTATGAGCTGGGACCACTTACATGGCTTTTGGGCTGCAGATA 493
 |||||
 208 eLysAsnLeuValLeuLysIleIleSerGlySerPheProProValSer 224
 |||||
 492 TGAACACCCCTGTGACTGAAATATATCTGATCTTTTCCACCTGTCT 443
 |||||
 225 LeuHisTyrSerTyrAspLeuAspArgSerLeuValSerGlnLeuPheLysAr 241
 |||||
 442 TTGCATATATTCCTATGATCTCCGACATTGGGTGTCTCAATTTTAAAG 393
 |||||
 241 gAspProAlaSerGlyGlnAsnSerIleLeuGluLysGlyPheI 258
 |||||
 392 AAATCTTAGGATAGACCATCAGTCATCTCATATTGGAGAAAGTTTAA 343
 |||||
 258 LeuAlaLysArgIleGluLysPheLeuSerProGlnLeuIleAlaGlu 274
 |||||
 342 TAGCCAAACGCATTGAAAGTTCTCTCTCCACGCTTATTCAGAGAA 293
 |||||
 275 PheCysLeuLysThrPheSerLysPheGlySerGlnProIleProAlaLys 291
 |||||
 292 TTTGTCTAAACATTTTCCAGATTGGATCCACACCTATACAGCTAA 243
 |||||
 291 sArgProAlaSerGlyGlnAsnSerIleSerValMetProAlaGlnLysI 308
 |||||
 242 AAGACCACTTCAGACAAACCTCGANTTCTGTATGCCCTGCAGAA 193
 |||||
 308 IeThrLysProAlaAlaLysTyrGlyIleProLeuAlaTyrLysLysTyr 324
 |||||
 192 TTACAAACCTGCCGCTAAATATGGAATACCTTTACATATTAAGAAAT 143
 |||||
 325 GlyAspLysLysLeuHisGluLysLysProLeuGlnLysHisLysGlnAl 341
 |||||
 142 GGAGATTAATAAATTATACAGAAAGAAACCACTGCAAAACATTAACAGGC 93
 |||||
 341 aHisGlnThrProGluLysArgValAsnThrGlyGluGluAlaArgLysI 358
 |||||
 92 CCATCAAACTCCAGAGAAAGAGATGATCTGAGAAAGAAAGAGGAA 43
 |||||
 358 IeSerGluGluAlaAlaArgLysArgArgLeuGluIlePheI 371
 |||||
 42 TATCTGAGGAAGCAGCAAGAAAGAAAGCGCTGGAATTTATT 2
 |||||
 seq_name: gb_est1:AM976435
 seq_documentation_block:
 LOCUS AM976435 654 bp mRNA linear EST 02-JUN-2000
 DEFINITION EST38834 MAGE resequences, MAGN Homo sapiens cDNA, mRNA sequence.
 ACCESSION AM976435
 VERSION AM976435.1 GI:8167663
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 1 (bases 1 to 654)
 Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspar,R., Gay,C., Holt
 ,I.E., Saeed,A.I., Sharov,Y., Lee,N.H., Yeatman,T.J. and
 Quackenbush,J.
 Assessment of gene expression patterns in a model of colon tumor
 metastasis using a 19,200 element cDNA microarray
 JOURNAL Unpublished (2000)
 COMMENT Contact: John Quackenbush
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 3528
 Fax: 301 838 0208
 Email: johng@tigr.org
 Plate: 360
 Seq primer: Forward.
 Location/Qualifiers
 1. 654
 /organism="Homo sapiens"

```

/db_xref="taxon:9606"
/clone_lib="MAGE_resequences_MAGN"
/note="Vector: BluescriptSKm"
BASE COUNT      171 a      134 c      121 g      228 t
ORIGIN

alignment_scores:
    Quality: 1016.00      Length: 215
    Ratio: 4.861          Gaps: 0
    Percent Similarity: 97.209      Percent Identity: 95.349

alignment_block:
US-09-783-320-4 x AW976435/rev ..

Align seg 1/1 to reverse of: AW976435 from: 1 to: 654

157 ValGluLeuAlaArgThrCysIleGlyThrProTyrTyrLeuSerProG1 173
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
643 CTGAGAGAGGGGCCGACCTTGATGAGGGGCCCACTGCTGCTGCTGCA 594
173 uilecysgluasnlspoftryasnanslysserapilletppalaleu 190
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
593 AATCGTGTAACAACACT.TACATAATATAAAG.GACATTGGGCTCTGG 546
190 llycysvalleutrgluLeuCysThrleuIshisAlapheglualagly 206
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
545 GGTGTGCTCTTATGAGCTGTGACACTTAAACATCTTTGAAAGCTGG 496
207 SerMetLysAsnLeuValleuLysIleIleSerGlySerPheProProVa 223
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
495 AGTATGAAAAAACCCTGCTACTGACATGATATCTGATCTTTCACCTGT 446
223 lserleuIshySerTyrAspLeuArgSerLeuValSerGlnLeuPheL 240
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
445 GTCTTGCAATATTCCTATGATCTCCGACGTTGGTGTCTCAGTTATTTA 396
240 VSATASnProArgPargPargProSerValasnsertileuGlnLysGly 256
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
395 AAGAAATCTTAGGGATGACCATGACATCACTCATATTGGAGAAAGGT 346
257 PheIleAlaLysArgIleGluLysPheLeuSerProGlnLeuIleagl 273
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
345 TTATAGCCAAACGCAATGAAAGTTCTCTCTCCTCAGCTTATGAGAG 296
273 uGluPheCysLeuLysThrPheSerLysPheGlySerGlnProIlePro 290
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
295 AGAATTTGTCTTAAACATTTTTCGAACTTTCGATCAGACCTATACAG 246
290 lAlysArgProAlaSerGlyLysAsnSerIleSerValMetProAlaGln 306
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
245 CTAAAGACCCAGCTTCAGACAAACCTGATTTCTGTTATGCCCTGCTCAG 196
307 LysIleThrLysProAlaAlaLysTyrGlyIleProLeuAlaTyrLysLy 323
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
195 AAAATATCAAAAGCCTCCGCTAAATATGGAATACCTTATGCAATTAAGAA 146
323 sTYRGlyAspLysLysLeuHisGluLysLysProLeuGlnLysHisLysG 340
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
145 ATATGGAGATTAATAAATATACAGAAAGAAACCACTGCCAAAAACATTAAC 96
340 lAlaHisGlnThrProGluLysArgValasnthrGlyGluGluLysArg 356
95 AGGCCCATCAAACTCCAGAGAAAGAGTGAATATCTGAGAGAAAGAGG 46
357 LysIleSerGlnLysLysAlaAlaArgLysArgLeuGluPheIle 371
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
45 AAAATATCTGAGAGACAGCAAGAAAGAGAGCTGGATTTATT 1

seq_name: gb_est2:BF541163
seq_documentation_block:
LOCUS      BF541163              845 bp      mRNA      linear      EST 11-DEC-2000

```

```

DEFINITION      602068802P1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:406765 5',
ACCESSION      mRNA sequence.
VERSION      BF541163
AUTHORS      BF541163.1 GI:11628544
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      NIH_MGC http://mhc.nci.nih.gov/.
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE      Unpublished (1999)
JOURNAL      Contact: Robert Strausberg, Ph.D.
COMMENT      Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LICM907 row: d column: 02
High quality sequence stop: 672.
Location/Qualifiers
1. 845
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:406765"
/clone_lib="NIH_MGC_58"
/tissue_type="hyponephroma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Kidney; Vector: pDNR-LIB (Clontech); Site:1:
SfiI (ggcgctagcgc); Site:2: SfiI (ggcattatggc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
sequence: 5'-ATTCTAGAGGCGCGCGCGCATG-dt(30)BN-3'
adaptor (where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.35 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto CA)."
BASE COUNT      272 a      146 c      209 g      218 t
ORIGIN

alignment_scores:
    Quality: 1006.50      Length: 292
    Ratio: 4.142          Gaps: 9
    Percent Similarity: 83.219      Percent Identity: 78.082

alignment_block:
US-09-783-320-4 x BF541163 ..

Align seg 1/1 to: BF541163 from: 1 to: 845

913 GlnLeuSerAspIleHisIleGluProGlyThrAsnAspSerGlnHisSe 929
1 CAATTTAGTGCATTCATATAGAGCTTGGAACCAATGATTCACAGCAGCTC 50
929 rlycysaspyalaspyseryalginplogluprophepnehislysv 946
51 TAAATGATCTAGATAAGCTGTGCACCGGACCATTTTCCATAAGG 100
946 aValHisSerGlnHisLeuAsnLeuValProGlnValGlnSerValGln 962
101 TGGTTCATCTGAAACACTTGAAGCTTACCTCCCTCAAGTTCAATCAGTTCAG 150
963 CysSerProGluGluSerPheAlaPheArgSerHisSerHisLeuProP 979
151 TGTTCACCAAGAAATCCCTGGCATTTGCAATTCATCCCATTTACACAC 200

```


290 AlAtATgATGProAlaSerGlyGlnAsnSerIleSerValMetProAlaI 306
 248 GCTAAAGACCAAGCTTTCAGACAAACTCGATTCTGTATGCTCTCA 199
 306 nlystIlePrlYsProAlaIleAlaIleTyGlyIleProLeuAlaTyLysL 323
 198 GAAATATTAAGAGCTGCGCTAAATATGAAATCTTTAGCATATATAGA 149
 323 ystTyGlyAspLysLysLysLysLysLysLysLysLysLysLysLys 339
 148 AATATGAGATTAATAAATTACAGAAAGAACACACGCAAAACATATA 99
 340 GlnAlaHisGlnThrProGlnLysArgValAsnThrGlyGlnGluArg 356
 98 CAGGCCATCAAACTCCAGAGAGAGAGATCTGAGAGAGAGAAAGAG 49
 356 gLyAlleSerGluGlnAlaAlaArgLysArgArgLeuGluPheIle 371
 48 GAAATATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTTAT 3

seq_name: gb_est1:AA907081

seq_documentation_block:

LOCUS AA907081 477 bp mRNA linear EST 26-AUG-1998
 DEFINITION O103a06.s1 NCI_CGAP_Lus Homo sapiens cDNA clone IMAGE:1522354.3,
 similar to SW:NEK1_MOUSE P51954 SERINE/THREONINE-PROTEIN KINASE
 NEK1 ;, mRNA sequence.

ACCESSION

AA907081

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 477)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: egadsr1@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.S.E. web page: http://www-bio.lit.nih.gov/bdrp/image/image.html
 Insert Length: 461 Std Error: 0.00
 Seq primer: -40ml3 fwd. RT from Amersham
 High quality sequence stop: 373.
 Location/Qualifiers
 1..477

FEATURES

source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="1522354"
 /clone_id="NCI_CGAP_Lus"
 /tissue_type="carcinoid"
 /lab_host="DH10B"
 /note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
 modified polylinker; 1st strand cDNA was prepared from
 neuroendocrine lung carcinoid, and was then primed with a
 Not I - oligo(dT) primer. Double-stranded cDNA was ligated
 to Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT73 vector. Library is normalized. Library was
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT

124 a 90 c 89 g 174 t

alignment_scores:

Quality: 813.00 Length: 159
 Ratio: 5.113 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-783-320-4 x AA907081/rev ..

Align seg 1/1 to reverse of: AA907081 from: 1 to: 477

213 LeuYsLleIleSerGlySerPheProProValSerLeuHisTySerTy 229
 477 CTGAATATATATCTCGATCTTTTCACCTGTGCTTGCATATATCCTA 428
 229 rAspLeuArgSerLeuValSerGlnLeuPheLysArgAsnProArgSPA 246
 427 TGAATCTCGAGATTGTGTCTCATTTTAAAGAAATCTTAGGGATA 378
 246 rPProSerValAsnSerIleLeuGlnLysGlyPheIleAlaLysArgIle 262
 377 GACCATCAGTCAACTCCTCATTTTGAGAAAGATTTTATAGCCAAAGCAT 328
 263 GluLysPheLeuSerProGlnLeuIleAlaGlnGluPheCysLeuLysTh 279
 327 GAAAGATTCT 278
 279 rPheSerLysPheGlySerGlnProIleProAlaLysArgProAlaSerG 296
 277 ATTTTCAGAGTTTGATCAGACGCTTACCACTTACCACTTAAAGACGCTT 228
 296 lYglnAsnSerIleSerValMetProAlaGlnLysLleThrLysProAla 312
 227 GACAAACTCGATTTCGTATCTCCCTGCTCAGAAATTAACAAAGGCTGCC 178
 313 AlAtATgATGProAlaIleAlaIleTyGlyLysTyGlyAspLysLysL 329
 177 GCTAATATGGAATACCTTATGCAATATGAATAATATGAGATAAAATTT 128
 329 nlystIlePrlYsProLeuGlnLysLysLysLysLysLysLysLysLys 346
 127 ACACGAAAGAAACACACGCAAAACATTAACAGGCCCTCAACTCCAG 78
 346 lAlaArgValAsnThrGlyGlnLysArgArgLysLleSerGlnGluAla 362
 77 AGAAGAGCTGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 28
 363 AlAtATgATGProAlaIleAlaIleTyGlyLysTyGlyAspLysLysL 371
 27 GCAGAGAGAGAGAGAGGCTGCAATTTATT 1

seq_name: gb_est1:AL036359

seq_documentation_block:

LOCUS AL036359 544 bp mRNA linear EST 29-FEB-2000
 DEFINITION DKFZP564C153.F1 564 (synonym: hfb2) Homo sapiens cDNA clone
 DKFZP564C153.5', mRNA sequence.

ACCESSION

AL036359

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 544)
 Mammut,R., Haubner,D., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
 EST (Mammut, et al.)
 Unpublished (1999)
 Contact: Mammut R
 MIPS
 Am Klopersplitz 18a D-82152 Martinsried, Germany
 This is the 5' sequence of the clone insert
 clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

THIS PAGE BLANK (USPTO)


```

117 shhValhAspArgLysIleuNHsArgSpIleLysSerGlnAsn 134
|||||
408 ACATGTACATGATGAAAAATTTCTTCAGACATTAATCTCAGACA 457
134 lPheLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAla 150
|||||
458 TATTTTAACTAAAGATGGAACACTACACTGGAGATTTTGGAAATTCG 507
151 ArgValIleuAsnSerThrValGlnLeuAlaArgPheCysIleGlyThr 167
|||||
508 ACAGTCTTAAATAGACTGAGACTGGCTGCAACTTGCAATGAGGACCC 557
167 cTtYrTrLeuSerProGluIleCysGluAsnLysProTyrAsnAsnLys 184
|||||
558 ATACTACTGTGACCTGAAATCTGGAACAACCTTACAAATATATAAA 607
184 eRAspIleTrpAlaLeuGlyCysValLeuTyrGlnLeuCysThrLeuLys 200
|||||
608 GTGACATTTGGCTCTGGGGTGTCTCTTAATGAGCTGTACACTTAA 657
201 hIsAlaPheGluAlaGlySerMetLysAsnLeuValLeuLysIleIle 217
|||||
658 CATGCTTTTGAAGCTGCGAGTATGAAAAACCTGCTACTGAAATATATC 707
217 rGlySerPheProProValSerLeuNHsTyrSerTyrAspLeuArgSer 234
|||||
708 TGGATCTTTTCCACCTGTCTCTTTCATTTATCTATGATCTCCGAGTT 757
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
|||||
758 TGGTGTCTCAGTTATTTAAAGAAATCTAGGAGTACCATCATGCTAAC 807
251 SerIleLeuGluLysGlyPheIleAlaLysArgIleGluLysPheLeu 267
|||||
808 TCCATATTGGAGAAAGTTTATAGCCAAAGCAATGCAAAAGTTCTCTC 857
267 rProGlnLeuIleAlaGluGluPheCysLeuLysThrPheSerLysPhe 284
|||||
858 TCCTCAGCTTATTCAGAAAGATTTTGTAAAAACATTTTCGAAAGTTTG 907
284 lYSerGlnProIleProAlaLysArgProAlaSerGlnLysAsnSerIle 300
|||||
908 GATCACACGCTATACCACTAAAGACCACTTCAGAGCAAAACCTCAT 957
301 SerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyrGly 317
|||||
958 TCTGTATGCGCTGCAGAAAAATTACAAAGCCTCCGCTAAATATGAA 1007
317 eProLeuAlaTyrLysLysTyrGlyAspLysLysLeuNHsGlnLysLys 334
|||||
1008 ACCTTTAGCATATTAAGAAATATGGAGATAAAATTTACAGAAAGAAC 1057
334 rGlnLeuGlnLysHisLysGlnAlaHisGlnThrProGluLysArgValAsn 350
|||||
1058 CACTCGAAACATTAACAGGCCCATCAACTCCAGAGAGAGAGCTGAAT 1107
351 ThrGlyGluGluArgArgLysIleSerGlnGluAlaAlaArgLysArg 367
|||||
1108 ACTGGAGAGAAAGAGGAAAAATCTCAGGAGACCAAGAGAAAGAAAG 1157
367 gLeuGluPheIleGluLysGluLysGlnLysAspGlnIleIleSerL 384
|||||
1158 GCTGGAAATTTATTTGAAAAAGATAAGAACGGTAGATCAGATTAAT 1207
384 euMetLysAlaGluGlnMetLysArgGlnGluLysGluArgLeuGluArg 400
|||||
1208 TAATGAGGCTGAACAATGAAAAAGCCAAAGCAAGAAAGTTGGAAGA 1257
401 lIleAsnArgAlaArgGluGlnGlyTyrPArgsnValLeuSerAlaGly 417
|||||
1258 ATTAATAGGCGCAAGCAACAAGATGGAATAATGCTTAAGTCTGTG 1307

```

```

417 ySerGlyLysValLysAlaProPheLeuGlySerGlyGly.....Thr 432
|||||
1308 AAGGTGAGAGTAAGTAGGCATTTTATATCCAAATATGTTATCTACCA 1357
432 lAlaProSerSer 436
|||||
1358 TTTTCCCTCCAGT 1371
seq_name: /cgn2_6/ptodata/1/ina/6B_COMB.seq:us-09-420-915-15
seq_documentation_block:
? Sequence 15 Application US/09420915
? Patent No 6264947
? GENERAL INFORMATION:
? APPLICANT: Bandman, Olga
? APPLICANT: Tang, Y. Tom
? APPLICANT: Hillman, Jennifer L.
? APPLICANT: Yue, Henry
? APPLICANT: Guebler, Karl J.
? APPLICANT: Corley, Neil C.
? APPLICANT: Gorgone, Gina
? APPLICANT: Azimzal, Yalda
? APPLICANT: Lu, Alina
? TITLE OF INVENTION: Protein Kinase Homologs
? FILE REFERENCE: PF-0614 US
? CURRENT APPLICATION NUMBER: US/09/420,915
? EARLIER FILING DATE: 1999-10-20
? EARLIER APPLICATION NUMBER: US 09/173,581
? NUMBER OF SEQ ID NOS: 18
? SOFTWARE: PERL Program
? SEQ ID NO 15
? LENGTH: 1846
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE: -
? OTHER INFORMATION: 1567782
? US-09-420-915-15
alignment_scores:
? Quality: 1959.00 Length: 438
? Ratio: 4.922 Gaps: 2
? Percent Similarity: 90.868 Percent Identity: 89.269
alignment_block:
US-09-783-320-4 x US-09-420-915-15
Align seg 1/1 to: US-09-420-915-15 from: 1 to: 1846
1 MetGluLysTyrValArgLeuGlnLysIleGlyGluGlySerPheGly 17
|||||
154 ATGGAGAAAGTATGTAGTACACAGAAAGATTGGAGAAAGTTCTTTGAAA 203
17 sAlaIleLeuValLysSerThrGluAspGlyArgGlnTyrValIleLys 34
|||||
204 AGCCATTCCTGTATATCTACAGAGATGGCAGACGATGTATCAAG 253
34 lIleAsnIleSerArgMetSerSerLysGluArgGlnLysSerArg 50
|||||
254 AAATTAACATCTCAAGAAATGTCAGTAAAGAAAGAGAAAGAAATCAAG 303
51 GluValAlaValLeuAlaAsnMetLysHisProAsnIleValGlnTyr 67
|||||
304 GAACTTCAGTATTTGGCAACATGAAGACATCCAAATATTTGCTCACTA 353
67 gGluSerPheGlnGluLysnGlySerLeuTyrIleValMetAspTyrCys 84
|||||
354 AGAATCATTTGAA..... 366
84 lGluGlyLysAspLeuPheLysArgIleAsnAlaGlnLysIleValLeu 100
|||||
366 ..... 366

```

```

101 GINGLUAAPGlnlleuAspTrpPheValGlnlleCysLeuAlaLeuLy 117
    |||||
367 .....GGAAATTTGGACTGGTTGTACAGATATGTTGGCCCTGAA 407
117 SHSVAHLSASPARGLysIleLeuHLSARqAspIleLysSerGlnAsnI 134
    |||||
408 ACATGTACATGATGAAAAATTCCTCAGCGACATTAATCTCAGAACAA 457
134 LephLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAla 150
    |||||
458 TATTTTAACATAAGATGAGACAGACAACTTGAGATTTGGAAATGCT 507
151 ArgValLeuAsnSerThrValGlnLeuAlaArgThrCysIleGlyThrPr 167
    |||||
508 AGAGTTCTTAATAGTACTGTAGACTGGCTGCACTTCATCAGGAGCC 557
167 GYTYTYLLeuSerProGlnlleCysGlnAsnLysProGlyTrpAsnLys 184
    |||||
558 ATACTACTTGTCCACGTGAATCTGTGAATAACAACCTTACATAATATA 607
184 eAspIleTrpAlaLeuGlyCysValLeuTYrGlnLeuCysThrLeuLys 200
    |||||
608 GTGACATTTGGCTCTGGGGTGTCTTATGAGCTGTGTACACTTAA 657
201 HisAlaPheGlnAlaGlySerMetLysAsnLeuValLeuLysIleIle 217
    |||||
658 CATGCTTTGAGCTGCGAGTATGAAAAACCTGGTACTGAGATATAATTC 707
217 rGlySerPheProProValSerLeuHLSYrSerTYrAspLeuArgSerI 234
    |||||
708 TGGATCTTTTCCACCTGTCTTGCATATTCCTATGATCTCCCACTT 757
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
    |||||
758 TGGTGTCTCACTTATTTAAAGAAATCCTAGAGATGACCATCACTAAC 807
251 SerIleLeuGlnLysGlyPheIleAlaLysArgIleGlnLysPheLys 267
    |||||
808 TCCATATTTGAGAAAGTTTATATGCCAAACCGCATTTGAAAGTTCTCTC 857
267 ProGlnLeuIleAlaGlnLysPheCysLeuLysThrPheSerLysPheG 284
    |||||
858 TCCCTAGCTATTTGCAGAAATTTTGTCTAAACATTTTGGAAATTTG 907
284 LysSerGlnProIleProAlaLysArgProAlaSerGlyGlnAsnSerIle 300
    |||||
908 GATCACAGCTATACAGCTAAAGACAGCTTCAGGACAAAACTGAT 957
301 SerValMetProAlaGlnLysIleThrLysProAlaAlaLysTYrGly 317
    |||||
958 TCTGTATGCCCTGCAGAAAAATTACAAAGCTGCCCTTAATATGAAAT 1007
317 eProLeuAlaTYrLysLysTYrGlyAspLysLysLeuHLSGlnLysSer 334
    |||||
1008 ACCTTTACGATATAAGAAATATGAGATTAATAAATTTACAGAAAGAAC 1057
334 rGlnLeuGlnLysHLSLysGlnAlaHLSGlnThrProGlnLysArgValAsn 350
    |||||
1058 CACTGCAGAAAACATAACAGGCCCTCAAACTCCAGAGAGAGTGAAT 1107
351 ThrGlyGlnGluArgArgLysIleSerGlnGlnAlaAlaArgLysArg 367
    |||||
1108 ACTGGAAGAAAGAGGAGAAATATCTGAGAGACAGACAGAAAGAGAAAG 1157
367 GlnGluPheIleGlnLysGlnLysGlnLysAspGlnIleIleSerL 384
    |||||
1158 GCTGGAAATTTATGAAAAAGATTAAGAACCGTAGGATAGATTTTATG 1207
384 euMetLysAlaGlnGlnMetLysArgGlnGlnLysGlnLysArgLeuGlnArg 400
    |||||
1208 TAATGAAGCTGAACAATGAAGAAAGCAAGCAAGAAAGTTGAAAGA 1257

```

```

401 ILeasArgAlaArgGlnGlnGlyTrpArgAsnValLeuSerAlaGly 417
    |||||
1258 ATTAATAGGCCGAGCAAGATGAGAAATGCTCTAGTCTGTG 1307
417 YserGlyGlnValLysAlaProPheLeuGlySerGlyGly.....ThrI 432
    |||||
1308 NAGTGTGAAGTAAGTAGGATTTTATACCAATATGTTATACACCA 1357
432 LeAlaProSer 436
    |||||
1358 TTTTCCCTCCACT 1371

seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:us-08-870-529-1
seq_documentation_block:
; Sequence 1, Application US/08870529
; Patent No. 608057
;
; GENERAL INFORMATION:
; APPLICANT: Sims, John E.
; APPLICANT: Virca, G. Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: IL-1/TNF- $\alpha$ -ACTIVATED KINASE (ITAK),
; TITLE OF INVENTION: AND METHODS OF MAKING AND USING THE SAME
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; City: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/870,529
; FILING DATE: 06-JUN-1997
; CLASSIFICATION: 435
;
; ATTORNEY/AGENT INFORMATION:
; NAME: McMaisters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.418
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2940 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2937
; US-08-870-529-1

alignment_scores:
Quality: 600.00 Length: 392
Ratio: 2.353 Gaps: 8
Percent Similarity: 65.051 Percent Identity: 33.673

alignment_block:
US-09-783-320-4 x US-08-870-529-1 ..
Align seg 1/1 to: US-08-870-529-1 from: 1 to: 2940

4 TYrValArgLeuGlnLysIleGlyGlnGlySerPheGlyLysAlaIle 20
|||||
114 TACATCCCATTCGCGCTCTGCGCGCGCGCTTCGGGAGACCGACGCT 203

```

20	uVallyserPhrghuAerGlyAuhgIntYValIleuysIuleIsnt	37
204	GRACSSCGCCACCGAGAGTACTACTCGTGTGTGGGAAGAACTGATT	253
37	IAsErAtgMeSerSerIySGuAargIuSerAAtgAGIuValAa	53
254	TCACCCGGCTGTCTAGAAAGAACTCGTATGCTTGAAGAAATAGTT	303
54	ValLeuAlaAsnMeLyshNzProAsnIleValGlnTyAArgIuSerPh	70
304	ATTCTGGCACTGCTCTACACAGCAACATATATGCGTACTCAATCACT	353
70	eGIuIuAsnGlySerIeTyuTyIleValMeAspTyCysGIuIyGlyA	87
354	CATGGACAAATACCAGCTGCTGATGTGACTCGAATATGTGTAAGAGGA	403
87	gPhLeuPhelysArgIleAsnIaGlnIySGIyValIleuPhGlnIuAsp	103
404	ACCTGTATGACAAATTCCTTGCTGTCAAGAGCAAACTTTTGAGAGAAAG	453
104	GlnIleleuAspTrpPhleValGlnIleCysleuAlaIeulysNzIshValnI	120
454	ATGGGAGGTGTGGTACCTATTCTCAGATTGTTCACAGAGGACTGCATCCA	503
120	sAsparIyGlyIleIeuNzAsArgAspIleIySerGlnAsnIlePhleuT	137
504	TAAACCTGGAAATCCTTCATAGAGATATGAACATATTAATTTTTCGTA	553
137	hLyAsArdIyThArgIleuGlnIleuGlyAspPhegIyIleAlaArgValIeu	153
554	CSAAGGCAAACTGTAAAACTGGAGATATGCGCTGACAAAGAACTT	603
154	AsnSerThValGlnIleAlaArgPhrGlySylleGIyThProTyGlyT	170
604	ATTCTGAATATTCATGCGTCAAGACGCTGTGGGAACCCATATTCAT	653
170	userProGlnIleCysGlnIuAsnIyProTyAsnAsnIySerAspIleT	187
654	GTCTCCAGAGCTGTCTCAAGAGGATGAATCAACATTTCAAGTGTATATCT	703
187	rPalAleuGIyCysValIeutyGlnIeuCysThIleuIyNzAlaArhe	203
704	GGCGAGTGGCGCTCATATTTTGAAACGCTTACCCTTAAAGAGACGTTT	753
204	GlnAlaGlySerMeLyAsnIleuValIeulysIleIeSerGly....	218
754	GATGCTACAAACCCCTTAACCTGTGTGAAAGATCGTGCAGAGATTCG	803
219	serPherProValAsnIleuNzIyTySerTyAspIleuIyGSeIeuY	235
804	GGCCATGGAATGTACTAGCCACATCTCTTGGAATTTGATCCAAAGG	853
235	AlSerGlnIleuPhelysArgAsnProArgAspArgProSerValAsnSer	251
854	TTTCATTCGGCTTGAACGAGCTCTGTGACAGACGCTACAGTGCAGATGA	903
252	IleleuGIuLyGIyPhelIleAlaIyArgIle..GlnIySpheluSerP	268
904	CTTCTAGAGATGGCCCTTCTCTCAGGAAAGCGAGAGAGAGATGAGAGAA	953
268	roGlnIleuIleAlaGIuIuPhelysIyIyIyIyIyIyIyIyIyIyIyIyIy	284
954	AGTCACCTC.....TCTTAATGACACTCAAAAGAGACCA	988
285	serGlnProIleProAlaIyAsnIyProAlaSer.....	295
989	GGTCAGACGCTGTGACTAGACACCCATTCCTGTATGTAACATCAGAAC	1038
296G	296

[illegible]

```

296 .....G 296
1158 AGCAAGCTCATGTTGGGGGTGGGAAATTCACCCCCAGAAACTGA 1207
296 1yglinserslleserValmetProAlaGlnIleEthylsProAla 312
1208 TGTTATCAACAGTGGCTGTATAGTCCCGGAGGTGTGTG..... 1245
313 AlAlaLysTyrGlyIleProLeuAlaLysLysTyrGlyAspLysLysL 329
1246 .....CAGCAATACCCACTTTCGTGTGTGTACAGTGGAGAGCAACTG 1288
329 euHisGlyLysProLeuGlnLysHisLysGlnAlaHisLysLThrPro 345
1290 TACAC.....TTGGGTGAACTATGCAAGGAGGCACCTAAACTCCA 1327
346 .....GluLysArgValAlaThrGlyLysGluLys 355
1328 TGTGTACGTGGGCCATGTGAACAAAGCTCTCTATGCACAGCCAAAGCATG 1377
355 gArgLysIleSerGlnLysAla 362
1378 TGGAAATCTTCAGAGCCAAAC 1399
seq_name: /cgn2_6/prodata/1/lna/6B_COMB.seq:US-09-393-569-1

```

```

seq-documentation block:
Sequence 1, Application US/09393569
Patent No. 627979
GENERAL INFORMATION:
APPLICANT: BINGHAM, SHARON
APPLICANT: CASSE, PARRICK
APPLICANT: LARSON, SALLY NEALE
APPLICANT: NEMTON, RICHARD ANTHONY
APPLICANT: PIERCE, OLIVER LARS
APPLICANT: RAUSCH, PRAVIN
APPLICANT: REITH, ALASTAIR DAVID
APPLICANT: SANGER, AGARETH JOHN
FILE OF INVENTION: NEW USE
FILE REFERENCE: P92261
CURRENT APPLICATION NUMBER: US/09/393,569
CURRENT FILING DATE: 1999-09-10
EARLIER APPLICATION NUMBER: GB 9907261.3
EARLIER FILING DATE: 1999-03-29
EARLIER APPLICATION NUMBER: GB 9819779.1
EARLIER FILING DATE: 1998-03-10
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 4181
TYPE: DNA
ORGANISM: HOMO SAPIENS
US-09-393-569-1

alignment_scores:
Quality: 493.00 Length: 1242
Ratio: 0.824 Gaps: 61
Percent Similarity: 48.148 Percent Identity: 22.866

alignment_block:
US-09-783-320-4 x US-09-393-569-1 ..

Align seq 1/1 to: US-09-393-569-1 from: 1 to: 4181

7 leuGlnysIleGlyGluGlySerPheGlyLysAlaIleLeuValLysSe 23
:.....:.....:.....:
180 GTGGACCTTGTCGAATGACATACACGGGCAAGTTATTAAGCGTCCTA 22
:.....:.....:.....:
23 rThrGluAspGlyArgGlnTyrValIleLeuGluIleAsnIleSerArgm 40
:.....:.....:.....:

```

230 TGTCAAAAGGGCCAGCTTTCAGCCATCAAGGTTATGATGTC..... 272
40 eTserSerLysGluArgGluGluSerArgGluValAlaValLeuAla 56
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
273 ..ACGGGGCATGAAAGAGAAATCAAAACAGAAATTAACATGTTGAAG 320
57 AsnMetLys...HisProAsnIleValGlnTyrArgGluSerPheGluGln 72
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
321 AAATATTCTCATCAACCGGAATATTGCTACATACATGCTGCTTTTTCAA 370
72 u.....AsnGlySerLeuTyrIleValMetAspTyrC 83
:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
371 AAAGAACCCACCAGCATGATGACCAACTTGTGTTGATGATGAGTTT 420
83 ySgIuGlyGlyAspLeuPheLysArgIleAsnAlaGlnLysGlyValLeu 99
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
421 GTGGTGGCTGGCTGTGTACCCGACCTGATCAAGAACACAAAGGTTAACAG 470
100 PheGlnGluAspGlnIleLeuAspTyrPheValGlnIleCys..... 113
471 TTGAAGAG.....GAGTGGATTGCAATACATCTGCAGGGAAAT 508
114 ...LeuAlaLeuLysHisValHisAspArgLysIleLeuHisArgAspI 129
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
509 CTTAGCGGGGCTGAGTCACTGCACACAGCATAAAGTATCATCGAGAT 558
129 IeLysSerGlnAsnIlePheLeuThrLysAspGlyThrValGlnLeuGly 145
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
559 TTAAGGGCAAAATGTCTTCTGACTGAATAAGCAAGTAAGTAAGT 608
146 AspPheGlyIleAlaArgValLeuAsnSerThrValGluLeuAlaArgH 162
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
609 GACTTGGAGTCAAGTCACTGATGCAACAGTGGGAGAGAGATAC 658
162 fCysIleLeuThrProTyrTyrLeuSerProGluIle.....CysGlu 177
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
659 TTTCATGGAACTCCCTACTGTGAGCAACAGATTAATGCTCTGTGATG 708
177 snLysPro.....TyrAsnAsnLysSerAspIleThrAlaLeuGly 190
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
709 AAAACCCAGATGCCACATATGATTTCAGAGATGACTTGTGTTGGGT 758
191 CysValLeuTyrGlnLeuCysThrLeuLysHisAlaPheGluAlaGly 207
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
759 ATCACCSCCATTTGAATGCA..... 779
207 rMetLysAsnLeuValLeuLysIleIleSerGlySerPheProValS 224
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
780GAAAGTGTCTCCCTCTGTGTG 801
224 eTLeuHisTyrSerTyrAspLeuArgSerLeuValSerGlnLeuPheLys 240
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
802 ACATGCAC.....CCCATGAGAGCTCTCTTC.....CTCATCCSC 836
241 ArgAsnProAlaArgAspArgProSerValAsnSerIleLeuGlnLysGly 257
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
837 CGGAATCCAGCGCTGG.....CTGAAGTCTTAAGA 868
257 eIleAlaLysArgIleGlnLysPheLeuSerProGlnLeuIleAlaGln 274
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
869 GTGGCAAAATAATCCAGTCAATTATTGAGAGCTGTTGTAATAAATC 918
274 IuPheCysLeuLeuThrPheSerLysPheGlySerGlnProIleProAla 290
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
919 AC.....AGC 923
291 LysArgProAlaSerGlyLys.....AsnSerIleSerValMet 303
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
924 CAGCGACGACGACAGAAATGATGAGACATTCATTTAGACAGACCA 973
303 rProAlaGlnLysIleThrLysProAlaAlaLysTyrGlyIleProLeu 320
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
974 ACCTAATAGAGGACAGTCCCATTCATTCAGTCAAGGACCATATTGATAGAA 1023

320 IatTyrLysTyrGlyAspLysLysLeuHisGluLysLysProLeuGln 336
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1024 CAAAGAAAGAGGACGAG.....GAAAAAGATCAGACAGAG 1058
337 LysHisLysGlnAlaHisGlnThrProGluLysArgValAsnThrGlyGln 353
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1059 TATGATAGACAGTGAAGTGAAGAGAGAGAGAGATGACTCGAGAGA 1108
353 uGlu.....ArgArg.... 356
1109 GCCCAGCTGCATCTGAATCTGCAGAGGAGTGCAGCCTGCGAGAGAT 1158
357LysIleSerGlnGlnAlaAlaArgLys 365
1159 TTCTGAGCTGCAGCTGSCCAACAGAGCCTTCTGAGGCTCTAGAGAG 1208
366 ArgArgLeuGlnIuPheIleGluLysGlyLysGln...LysAspGlnI 381
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1209 CAGCAGCTGAGCAGCAGCAGCGGAGATGAGAGCAGCAAGCCGACACT 1258
381 eIleSerLeuMetLysAlaGlnGlnMetLysArgGlnLysGlyArgL 398
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1259 GCTG.....GCCAGCGCTCAGAAAGCCTATCGAGGAGCAGAAAG 1296
398 euGluArgIleAsnArgAlaArgGluGlnGlyTyrArgAsnValLeuSer 414
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1297 AGCAGAGG...CGCGGCTGAGAGAGCAGCAAAAGCGA..... 1331
415 AlaGlyLysSerGlyLysValLysAlaProPheLeuGlnSerGlyLys 431
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1332GAGAGGAGCTCCCGAAGCAG..... 1352
431 rIleAlaProSerSerPheSerSerArgGlyLysIleTyrGlnHisTyrHis 448
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1353CAGAGAGGAGCAGCAGCGGCACTATGAG. 1382
448 IatIleAspGlnMetGlnGlnIleArgAlaGluAspAsnGluAlaLys 464
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1383GAGCAGATGCGCGGAGAGAG...GAGAGGAGGCTCGGAG 1421
465 TrpLysArgGlnIleTyrGluArgGlyLeuProGluArgGlnLys.... 479
1422 CATGACAGGAAATGATCAAGCAGCAGTGTAGAGAGAGCAGACAGCT 1471
480GlyIleLeuAlaValGluArgAla..... 487
1472 AGAGATCTGCAGCAGACGACTACTGATGACACAGCTCTATGGAAT 1521
488LysGlnValGluGluPheLeuGlnArgLysAlaIleMet 501
1522 ATAAAGCCAAACAAATTTGGAAGAA.....CAGACAGCAAGCAAAAGACTG 1565
502 GlnAsnLysAlaArgAlaGln...GlnHisMetValTyrLeuAlaArgLe 517
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1566 CAGAGCGAGCTAAAGCAAGAAAGAGACTACTAGTTCCCTTCAGACCA 1615
517 uArgGlnIleArgLeuGlnAsnPheAsnGluArgGlnGlnIleLysAla 534
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1616 GCGGCAAG.....GAGCAGAGGCTGTG..... 1637
534 yLysLeuArgGlyGlnLysGlnAlaAsnHis...SerGlnGlyGln 549
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1638GAGAGAGAGCCCTGATCAATTAAGAAAGAAATGAGT 1676
550 GlySerGlnGluLysAspMetArgArgLysIleGlnSerLeuLysAl 566
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1677 CTTAGTGAAG..... 1685
566 aHisAlaAsnAlaArgAlaValLeuLysGlnGlnLeuGlnArgLys 583
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1686AAGCCAGATGGGCCAAAGAGAGTAAAGAAAGCGGTCA 1722

583 rglYsgluAlatYrgluArgluLysValTrpGluGluHisLeuVal 599
|||
1723 GG.....CTCAACGGGAAAGTTCCCTCCATGCGCTCACAAAGGT 1763
|||
600 AlAlYsgluYalLysSerSerProLeuGluHisSgl 616
|||
1764 GCCACAGATATCTGATCCCAACCTGCCCCAAGGTGAGTCTTCAG 1813
|||
616 urhnglyGly.....SerProserL 623
|||
1814 CATTAGTGGAGTTACACCTGCTGCAACCCCCCATGCTCAGACCAAGTCG 1863
|||
623 YsgluGluMetArgSerValHisSerValHisSerAlaLeuLysGluVal 639
|||
1864 ATCCCGCATGCCATCTGGTAGCTGTAAATCCAG..... 1901
|||
640 GLyValAspSerSerLeuThrAspThrArgLutHisSerGluGluMet 656
|||
1902GGACCTGCTTGACCGCTCCAGTCCAGTGCACGACGACCCAC 1945
|||
656 nlygThrAsnAsnAlaHisSerSerLysArgLutLeuArgArgLeuA 673
|||
1946 AAAG.....GGCTCTCTGGGTTTCAGGAGGCTCTG..... 1976
|||
673 snGluAsnLeuLysAlaGluGluAspGluLysGluMetGluAsnLeuSer 689
|||
1977AACGTGACCTCCACCGCGGTGAGATGCCACGCGAAGTCAGAT 2021
|||
690 AspThrPheGluLeuAsnValHisGluAspAlaLysGluHisGluLysGlu 706
|||
2022 CCCACCTCGGAA.....AATCTCTCTCTCCCCAC 2050
|||
706 uLySerValSerSerAspArgLysTrpGluAlaGluGluLeuVal 723
|||
2051 TGGCATGTAAGTTGACCGAAGCTCTGTGTTACGACGAGAGAGAC 2100
|||
723 AlLLePro...LeuAspGluLeuThrLeuAspThrSerPheSerThrThr 738
|||
2101 TTCCACCAAGAGTGGCTCAAGAACACTTATATCCGACGATTAGCC 2150
|||
739 GluArgHisThrValGluValLLeuLysLeuPro.....AsnGlu 753
|||
2151 AGAAGAAATCTCTCGGATGCTGATGCTGTGGACCCAGACTAGATC 2200
|||
753 ySerProArgArgAlaTrpGluLysSerPro.....ThrAspS 766
|||
2201 TCACCCATCAGAGCA.....AGCACCTGATCTCCGAGAACTGAGC 2244
|||
766 erValLeuLysLLeuGluLysGluLysLeuGlu..... 777
|||
2245 CCAATCTTG.....GAGAGCCCTTGCAGAGGACGACGAGTGGC 2282
|||
777 777
|||
2283 AGTTCTCCAGCTCCAGACCCCTAGCTCCAGCCAGCTCCCAAGAGG 2332
|||
778LeuGluThrGluLeuLeuGluAsnThrThrLeuArgS 790
|||
2333 CTCACAGCTCGATCCACAGCAGAGTCAAGTGAAGCAGACAGAGTTCGAG 2382
|||
790 erGluLysSerProGluGluGluLysTrpLysProLeuLLeuThrGlyGlu 806
|||
2383 CCAACAGTACAGTCAAGAGTCCACTGCTCTCC..... 2417
|||
807 LysLysValGluCysLysSerHisGlu.....LeuAsnProSerAl 820
|||
2418CATGAGCTTGCACAGGTGAACCCGAGAA 2446
|||
820 AlLeuAlaAspSerProValGluThrLysSerProGluPheSerGluAlas 837
|||
2447 ATCCAGGACATTTACCGCGCCAGTGCACAGCTACTACAAAAAAGCTA 2496
|||
837 erProGluMetSerLeuLysLeuGluGluLysAsnLeuGluGluProAsp 853

2497 TAGATGAGATCTGACGCGATTAGCCAAAGAACTAAGAGAA..... 2537
|||
854 LeuGluThrGluLLeuLeuGluGluProSerGlyThrAsnLysAspGluSe 870
|||
2538 CTCGGATGTAAACAAACCCCAATGAAAGAGTGAATGATCTC 2587
|||
870 rLeuProCysThrLLeuThrAspValTrpLeSerGluGluLysGluThrL 887
|||
2588 CTCCTCC.....AGTGAAGGTGAGAAAGTA 2613
|||
887 YsgluThrGluSerAlaAspArgLLeuThrLLeuGluAsnGluValSer 903
|||
2614 GCGAG.....GAAGAGAG 2627
|||
904 GluAspGlyValSerSerThrValAsp.....GluLeuSerAspL 917
|||
2628 GAGATGTGAGAGACGAGACCATGATGGACAGATGGCTGACGAGAT 2677
|||
917 eHisLLe.....GluProGlyThrAsnAspSerGluHisS 929
|||
2678 ACCCAGACTGATACCAACAGGAGCTCCAGGACGACGAGAGTCAATG 2727
|||
929 erLysCysAspValAspLysSerValGluProGluProPheHisLys 945
|||
2728 TGGGAATGTGGGAGCATGCGCTGCGAGACTCT..... 2762
|||
946 ValValHisSerGluHisLeuAsnLeuValProGluValGluSerValGlu 962
|||
2763CATCGGACAGTTTC.....AGCGCGAG 2785
|||
962 nCysSerProGluGluSerPheAlaPheArgSerHisSerHisLeuProP 979
|||
2786 TATTTCAGAGAGAGAACCTTGATGATTAAGAGAGCTCT..... 2825
|||
979 rOlyAsnLysLysLysSerLeuLeuLLeuGlyLeuSerThrGlyLeu 995
|||
2826GGAGAGAGAGGATCT.....GCCGAC 2849
|||
996 PheAspAlaAsnAsnProLysMetLeuArgThrCysSerLeuProAspLe 1012
|||
2850 AGTGACAGCAATGGC.....TTTCTGCGCAATCAACTCTCCCTGACCT 2893
|||
1012 uSerLysLeuPheArgThrLeuMetAspValProThr.....ValGlyA 1027
|||
2894 GGTGCAGCAGAGCCATCTCCAGCTGGAACCCGACGAGGAGTGGGCGC 2943
|||
1027 spValArgLysAsnAsnLeuGluLysAspGluLLeuLysAspGluAsnLLe 1043
|||
2944 GCGTCTCAACCCATTCTCCAGGAGATG.....GACTCTGGGACT 2981
|||
1044 LysGluGlyProSerAspSerGluAspLLeuValPhe..... 1055
|||
2982 GAATATGGCAATGGGAGAGACCAAAAGCTCTTCCACCCCTTGTGGA 3031
|||
1056GluGluThrAspThrAspLeuGluGluLeuG 1066
|||
3032 CCCAGAGATATACAGACGCTCCCACTGATATAAGATAGAAGATGAGG 3081
|||
1066 LnaLaser.....MetGluGluLeuLeuArgGluGluPro 1077
|||
3082 AATCATCAGCGCGAGCTCTGTACTAGCAACTCTTATGACAGAAACAG 3131
|||
1078 GlyGluGluTrpSerGluGluGluGluSerValLeuLysAsnSerAspVa 1094
|||
3132 GCCAAATCATATGAGCAAGAAAGATTGCGTGTGTA..... 3167
|||
1094 LgluProThrLAsnGlyThrAspValAlaAspGluAspAspAsnPro. 1110
|||
3168AATGTAAACCAACCAACATTCGGGCTCTATACGACACACAG 3210
|||
1111 SerSerGluSerAlaLeuAsnGlu 1118

Wed May 15 14:14:01 2002

us-09-783-320-4.rml

Page 8

3211 AATCAGAAATACAGAAACGAT 3234

seq_name: /cgn2_6/ptodata/1/lna/6A_COMB.seq:us-09-221-235-3

seq_documentation_block:

Sequence 3, Application US/09221235

Patent No. 6043040

GENERAL INFORMATION:

APPLICANT: Acton, Susan

TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

FILE REFERENCE: MNI-050

CURRENT FILING DATE: 1998-12-28

EARLIER FILING DATE:

NUMBER OF SEQ ID NOS: 15

SOFTWARE: Patentln Ver. 2.0

SEQ ID NO 3

LENGTH: 906

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(906)

US-09-221-235-3

alignment_scores:

Quality: 485.50

Ratio: 2.503

Percent Similarity: 74.046

Percent Identity: 38.931

alignment_block:

US-09-783-320-4 x US-09-221-235-3

Align seg 1/1 to: US-09-221-235-3 from: 1 to: 906

```
8 GlnlyslleGlyGlySerPheGlyAlaIleValLysSerTh 24
|||||
112 AAGAAATATGTCGCGACAAATTTAGTAAGTTATAGCAGCGCTGCT 161
24 rGluAspGlyArgGlnTyValIleLysGluIleAsnIle...SerArgm 40
|||||
162 CTGGATGCAAGACACGCTGCTGATGATCAACAAAGATAGATCTTAAAG 261
40 etSerSerLysGluArgGlnGluSerArgArgGluValAlaValIleAla 56
|||||
212 TGGATGCAAGACACGCTGCTGATGATCAACAAAGATAGATCTTAAAG 261
57 AsnMetLysHisProAsnIleValGlnTyValArgGluSerPheGluLys 73
|||||
262 CAACCTCAACACCAAAATGTAATTAATATATGATCATTCATGTAAGA 311
73 ngLysSerLysTyValIleValMetAspTyCysGlnGlyLysPhe... 88
|||||
312 TAATGAACTTAACATAGTTTGGAACTAGCAATGCTGCACTTATCCA 361
89 ....PheLysArgIleAsnAlaGlnLysGlyValIleuPheGlnGluAsp 103
|||||
362 GAATGATCAACATTTTAAAGAAAGAGAG...CTAATTCCTGAAGA 408
104 GlnIleLeuAspTrpPheValGlnIleCysLeuAlaLeuLysHisValHt 120
|||||
409 ACTGTTGGAAAGATATTTGTCAGCTTTCAGTTCATGGAACACATGCA 458
120 SAAPATGlyLysIleLeuHisArgAspIleLysSerGlnAsnIlePheLeu 137
|||||
459 TTCTCAAGAGATCATGATAGATATTAACACCACTAATGCTGTCAATA 508
137 hrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAlaArgValLeu 153
|||||
509 CAGCCACTGGGCGTAAACTTGGAGATCTTGGCGCTTGGCCGCTTTTTC 558
```

seq_name: /cgn2_6/ptodata/1/lna/6A_COMB.seq:us-09-221-928-3

seq_documentation_block:

Sequence 3, Application US/09221928

Patent No. 6121030

GENERAL INFORMATION:

APPLICANT: Acton, Susan

TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

FILE REFERENCE: MNI-050

CURRENT FILING DATE: 1998-12-28

EARLIER FILING DATE:

NUMBER OF SEQ ID NOS: 15

SOFTWARE: Patentln Ver. 2.0

SEQ ID NO 3

LENGTH: 906

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(906)

US-09-221-928-3

alignment_scores:

Quality: 485.50

Ratio: 2.503

Percent Similarity: 74.046

Percent Identity: 38.931

alignment_block:

US-09-783-320-4 x US-09-221-928-3

Align seg 1/1 to: US-09-221-928-3 from: 1 to: 906

```
8 GlnlyslleGlyGlySerPheGlyAlaIleValLysSerTh 24
|||||
112 AAGAAATATGTCGCGACAAATTTAGTAAGTTATAGCAGCGCTGCT 161
24 rGluAspGlyArgGlnTyValIleLysGluIleAsnIle...SerArgm 40
|||||
162 CTGGATGCAAGACACGCTGCTGATGATCAACAAAGATAGATCTTAA 211
40 etSerSerLysGluArgGlnGluSerArgArgGluValAlaValIleAla 56
|||||
```

```

212 TGGATGCCAAAGCAGCTGCTGATTCATCAAGAAATAGATCTTTAG 261
57 AsnMetLysHisProAsnIleValGlnIleValArgLysSerPheGluAsp 73
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
262 CAACCTCAACACATCAAAATGTAATAAATATATGATCATCATTCATTAAGA 311
73 nglySerLeuTyrIleValMetAspTyrCysGluGlyIleAspLeu.... 88
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
312 TAATGAACATTAACATGTTTGGAACTAGACATGCTGCGACCTATCA 361
89 .....PheLysArgIleAsnAlaGlnLysGlyValLeuPheGlnIleAsp 408
362 GAATGATCAAGACATTTTAAAGCAAAAGAG...CTAATTCCTGAAGA 408
104 GlnIleLeuAspTyrPheValGlnIleCysLeuAlaIleValHisValH 120
409 ACTGTTCGAAGATTTTGTTCAGCTTTCAGTGCATTCGTAACACATCA 458
120 SASPARGLysIleLeuHisArgAspIleLysSerGlnAsnIlePheLeu 137
459 TTCTCGAAGAGTCATGCATAGAGATATATAAACACAGCTATGTGTTCA 508
137 hrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAlaArgValLeu 153
509 CAGCCACTGGGCTGCTAAACCTGGAGATCTGGGCTTGGCCGCTTTTC 558
154 AsnSerThrValGlnLeuAlaArgThrCysIleGlyThrProTyrTyrLe 170
559 AGCTCAAAAACACAGCTGCACATCTTTAGTTGGTACCCCTTATTACAT 608
170 userProGluIleCysGluAsnLysProTyrAsnAsnLysSerAspIle 187
609 GTCTCGAAGAGATACATCAATAATGATATCAACTTCAATCTGCACATC 658
187 rPalLeuGlyCysValLeuTyrGlnLeuCysThrLeuLysHisAlaPhe 203
659 GGTCTCTGGCTCTACTATATAGATGCTGCATTAACAAGTCCCTTTC 708
204 GluAlaGlySerMetLys.....AsnLeuValLeuLysIleLeuSer 218
709 TATGTGACAAATATGAAATTTACTCACTGTGTAAAGATATAGAACAGTG 758
218 ySerPheProProVal...SerLeuHisTyrSerTyrAspLeuArgSer 234
759 TGACTACCACTCTCTCTCTCATATCATCTATTCAGAAAGACTCGACAGT 808
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
809 TAGTTAATATGTGCATCAACCCAGATCCAGAGAACGACACAGACGTACCC 858
251 SerIleLeuGlnLysGlyPheIleAlaLysArgIle 262
859 TATGTTTATGAC.....GTACCAAAAGAGATG 885
seq_name: /cgn2_6/ptodata/1/lna/6a_comb.seq:us-09-221-527-3
seq_documentation_block:
; Sequence 3, Application US/09221527
; Patent No. 6146832
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNT-050
; CURRENT APPLICATION NUMBER: US/09/221,527
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 906
; TYPE: DNA
; ORGANISM: Homo sapiens

```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(906)
US-09-221-527-3

Alignment_scores:
  Quality: 485.50      Length: 262
  Ratio: 2.503        Gaps: 6
  Percent Similarity: 74.046      Percent Identity: 38.931

Alignment_block:
US-09-783-320-4 x US-09-221-527-3 ..

Align seg 1/1 to: US-09-221-527-3 from: 1 to: 906

      8 GlnLysIleGlyGluGlySerPheGlyLysAlaIleValLysSerThr 24
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
     112 AAGAAATTTGTCGCCGACATTTAGTGAAGTTTATAGACAGCCTGTCT 161
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
     24 rGluAspGlyArgGlnIleValIleLysGluIleAsnIle...SerArg 40
     162 CTGTGATGAGTACCACTGACCTTTAAAGAAAGTGCAGATATTTGATTAA 211
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
     40 eSerSerLysGluArgGluGluSerArgArgGluValAlaValLeuAla 56
     212 TGGATGCCAAAGCAGCTGCTGATTCATCAAAAGATATGATCTTCTTAG 261
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
     57 AsnMetLysHisProAsnIleValGlnIleValArgLysSerPheGluAsp 73
     262 CAACCTCAACATCCAAATGTAATAAATATATATGATCATTCATTAAGA 311
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
     73 nglySerLeuTyrIleValMetAspTyrCysGluGlyIleAspLeu.... 88
     312 TAATGAACATTAACATGTTTGGAACTAGACATGCTGCGACCTATCA 361
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
     89 .....PheLysArgIleAsnAlaGlnLysGlyValLeuPheGlnIleAsp 103
     362 GAATGATCAAGACATTTTAAAGCAAAAGAG...CTAATTCCTGAAGA 408
     104 GlnIleLeuAspTyrPheValGlnIleCysLeuAlaIleLysHisValH 120
     409 ACTGTTTGGAGTATTTGTTTCAAGCTTTCAGTGCATTCGTAACACATCA 458
     120 SASPARGLysIleLeuHisArgAspIleLysSerGlnAsnIlePheLeu 137
     459 TTCTCGAAGAGTCATGCATAGAGATATATAAACACAGCTATATGTGTTCA 508
     137 hrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAlaArgValLeu 153
     509 CAGCCACTGGGCTGCTAAACCTGGAGATCTGGGCTTGGCCGCTTTTC 558
     154 AsnSerThrValGlnLeuAlaArgThrCysIleGlyThrProTyrTyrLe 170
     559 AGCTCAAAAACACAGCTGCACATCTTTAGTTGGTACGCCCTTATTACAT 608
     170 userProGluIleCysGluAsnLysProTyrAsnAsnLysSerAspIle 187
     609 GTCTCGAAGAGATACATCAATAATGATATCAACTTCAATCTGCACATC 658
     187 rPalLeuGlyCysValLeuTyrGlnLeuCysThrLeuLysHisAlaPhe 203
     659 GGTCTCTGGCTCTACTATATAGATGCTGCTATCAAAAGCTCTTTC 708
     204 GluAlaGlySerMetLys.....AsnLeuValLeuLysIleLeuSer 218
     709 TATGTGACAAATATGAAATTTACTCACTGTGTAAAGATATAGAACAGTG 758
     218 ySerPheProProVal...SerLeuHisTyrSerTyrAspLeuArgSer 234
     759 TGACTACCACTCTCTCTCTCATATCATCTATTCAGAAAGACTCGACAGT 808
     234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250

```

Wed May 15 14:14:01 2002

us-09-783-320-4.rni

Page 10

```

      |||||.....:|||||.....:|||||.....:|||||.....:
809 TGGTTAAATATGCGCATCAACCCAGATCCAGAGAGACGACGAGCTCAC 858
      |||||.....:|||||.....:|||||.....:|||||.....:
251 SerIleLeuGluTylGlyPheIleIalalysArgIle 262
      :::::|||||.....:|||||.....:|||||.....:|||||.....:
859 TATGTTTATGAC.....:GTAGCAAAAGAGGAGTGTG 885

seq_name: /cgs2_6/ptcdat4/1/lna/6A_COMB.seq:US-09-221-236-3

```

```

seq=documentation_block:
  Sequence 3, Application US/09221236
  Patent No. 6146841
  GENERAL INFORMATION:
  APPLICANT: Acton, Susan
  TITLE OF INVENTION: NOVEL CSAK-P-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
  FILE REFERENCE: MNT-050
  CURRENT APPLICATION NUMBER: US/09/221,236
  CURRENT FILING DATE: 1998-12-28
  EARLIER APPLICATION NUMBER: 09/163,115
  EARLIER FILING DATE: 1998-09-29
  NUMBER OF SEQ ID NOS: 15
  SOFTWARE: PatcIntln Ver. 2.0
  SEQ ID NO 3
  LENGTH: 906
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE:
  NAME/KEY: CDS
  LOCATION: (1)..(906)
  US-09-221-236-3

```

alignment_scores:		.
Quality:	485.50	Length: 262
Ratio:	2.503	Gaps: 6
Percent Similarity:	74.046	Percent Identity: 38.931

```
alignment_block:
US-09-783-320-4 x US-09-221-236-3 . .
```

Align seg 1/1 to: US-09-221-236-3 from: 1 to: 906

```

8  GlnTylLeuGlnGlySerPheGlyLysAlaIleuValLysSerth 24
   ::::::::::::::::::::::::::::::::::::::::::::::::::::
112 AAGAAATTCGTCGGCGCAATTATGTAAGATTATAGACAGCCTGCT 161
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
24  rGluaspGlyArgGlnTyrValIleLysGluIleAsnIle...SerArgm 40
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
162 CTTCGATGGAGTACCGAGTACGCTTTAAAAAAGTCAGCATATTTATTTAA 211
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
40  etSerSerLysGluArgGlnGluSerArgArgGluValAlaValIleuAla 56
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
212 TCGATCCAAAGCAGCAGCTCGTATTCATCAACAAAGATAGATCTTTAAG 261
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
57  AsmMetLysIleProAsnIleValGlnTyrArgGluSerPheGluGluas 73
   CAACTCACACATCCAAATATTAATAAATATATGATCATTCATTGGAAGA 311
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
73  ngLysSerLeuTyrIleValMetAspTyrCysGluGluGlyAspLeu... 88
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
312 TAATGAACATAACATATGTTTGGAACTAGCAGATCTCGCACTATCCCA 361
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
89  ....PheLysArgIleAsnAlaGlnTylGlyValLeuPheGlnGluasp 103
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
362 GAATGATCAAGCATTTTAAAGACAAAGAGAG...CTAATTCCTGAAGA 408
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
104 GlnIleuAspTyrPheValGlnIleCysLeuAlaIleuLysSHSLVAlH 120
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
409 ACGTTCGGAAGATTTTCTTCACTTTCACAGCTTCACATGCATGGAACACATGCA 458
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
120 SAspArgLysIleLeuHisArgAspIleLysSerGlnAsnIlePheLeut 137
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
159 TTTCTCAAGAGCTCATCATAGAGATTAATAAACACAGCATTAATGTGTTCAATTA 508
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::

```

[illegible]

seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-221-416-3

```
seq_documentation_block:
; Sequence 3, Application US/09221416
```

```

: APPLICANT: Acton, Susan
: TITLE OF INVENTION: NOVEL GSApK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
: FILE REFERENCE: MNT-050
: CURRENT APPLICATION NUMBER: US/09/221,416
: CURRENT FILING DATE: 1998-12-28
: EARLIER APPLICATION NUMBER: 09/163,115
: EARLIER FILING DATE: 1998-09-23
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: Patenclin Ver. 2.0
: SEQ ID NO: 3
: LENGTH: 906
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(906)
: IS-09-221-416-3

```

```
alignment_scores:
  quality: 485.50
  ratio: 2.503
  percent_similarity: 74.046
  percent_identity: 38.931
```

```
alignment_block:
US-09-783-320-4 x US-09-221-416-3
```

Align seg 1/1 to: US-09-221-416-3 from: 1 to: 906

```

8  GlnLysIleGlyGlyGlnLysSerPheGlyLysAlaIleLeuValLysSerTh 24
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
112 AAGAAATATGGTCGCGCCACATTTATATGTGATGTATACAGACCGCTGTCT 161
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
24 rGluASPGLyArGInTyRValIleLysGluIleasnIle...SerArGm 40

```



```

112 AAGAAATTCGTCGCGACATTTAGTAACTTTATAGACAGCGCTGCT 161
24  rgluasgilyargintyvalileugsluileasnille...seratgm 40
162 CTGGATGGAGTACACAGTACCTTTAAAAAAGTCAGATTTTATTTAA 211
40  etSerSerlysgluargluserarggluvalalavallea 56
212 TGGATGCCAAGACGCTGCTGATGCATCAAGAAATATATCTTTAG 261
57  AsnMetlyshsProasnilevalgintyargluserPhegluas 73
262 CAACCTCAACCATCCAAATGTAATAATATATGCAATCATTCATTGA 311
73  nglySerleutyrllevalmetasptyrcysgluuglyaspleu... 88
312 TAATGAACTAAACATAGTTTGGACTAGACAGATGCTGCCGCTATCA 361
89  ....PheLySarGileasnaIaglnlysglyvalleuphegluas 103
362 GAATGATCAACATTTTAAAGACGAAAGAG...CTAATTCCTGAAGA 408
104 GlnIleuaspttrPhevalglnilecysleualaleuylshsValhI 120
409 ACTGTTTGAAGTATTTTGTTCACCTTTGCAGTCATTCGAAACATGA 458
120 sasparglysileleuhsaraplielysserGlnasnilepheleut 137
459 TTCTCGAAGAGTCATGCATAGATATTAACCAAGCTATGTCTCATTA 508
137 hrlYaspglyThrValglnleuglYasPheglyllealargValleu 153
509 CAGCCACTGGGGGTGTAACCTTGAGATCTGGGCTGGCGGTTTTC 558
154 AsnSerThrValglnleualargThrcysIleglyThrProtyrTyle 170
559 ACCTCAAAAACCCACAGCTGCACATCTTGTAGTGGTAGCGCTTATAC 608
170 userProgluilecysgluasnlysproyrasnasnlySeraspillet 187
609 GTCTCCAGAGCAATACATGAATAATGATACACTTAATCAATTCGATCT 658
187 rPalaleuIcysvalleuylrgluueucysThrlleuylshsIalphe 203
659 GGCTCTGGCTGCTACTATATAGATGAGGCTGCATTCACAAAGTCTTTC 708
204 GluAlaIglySerMetlys....AsnleuValleuylleleSerGI 218
709 TATGGGACAAATGATTTATACACTGTGTAAGAAAGATAGACAGTG 758
218 ySerPheProProval...SerleuHslyrSerTyAspIleargserI 234
759 TGACTACCACTCTTCTCTGATCATATTCAGAAAGACTCGACAGT 808
234 euValserGlnleupheylsarasnpProargAsparGproSerValasn 250
809 TACTTAATATGTGCATCAACCCAGATCCAGAAAGCAGACAGCTCACC 858
251 SerIleuGlnlysglyPheIlealaysargile 262
859 TATGTTTATGAC.....GTACCAAGAGGATG 885

seq_name: /cgn2.6/ptodata/1/ina/68-COMB.seq:US-09-593-553-3
seq_documentation_block:
; Sequence 3, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPE-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/593,553
; CURRENT FILING DATE: 2000-06-14

```

```

; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 906
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(906)
US-09-593-553-3

alignment_scores:
  Quality: 485.50      Length: 262
  Ratio: 2.503        Gaps: 6
  Percent Similarity: 74.046      Percent Identity: 38.931

alignment_block:
  US-09-783-320-4 x US-09-593-553-3

Align seg 1/1 to: US-09-593-553-3 from: 1 to: 906

      8 glnylsileglygluglyserPheglylYsAlaleuVallyserTh 24
      112 AAGAAATTCGTCGCGACATTTAGTAACTTTATAGACAGCGCTGCT 161
      24  rgluasgilyargintyvalileugsluileasnille...Seratgm 40
      162 CTGGATGGAGTACACAGTACCTTTAAAAAAGTCAGATTTTATTTAA 211
      40  etSerSerlysgluargluserarggluvalalavallea 56
      212 TGGATGCCAAGACGCTGCTGATGCATCAAGAAATATATCTTTAG 261
      57  AsnMetlyshsProasnilevalgintyargluserPhegluas 73
      262 CAACCTCAACCATCCAAATGTAATAATATATGCAATCATTCATTGA 311
      73  nglySerleutyrllevalmetasptyrcysgluuglyaspleu... 88
      312 TAATGAACTAAACATAGTTTGGACTAGACAGATGCTGCCGCTATCA 361
      89  ....PheLySarGileasnaIaglnlysglyvalleuphegluas 103
      362 GAATGATCAACATTTTAAAGACGAAAGAG...CTAATTCCTGAAGA 408
      104 GlnIleuaspttrPhevalglnilecysleualaleuylshsValhI 120
      409 ACTGTTTGAAGTATTTTGTTCACCTTTGCAGTCATTCGAAACATGA 458
      120 sasparglysileleuhsaraplielysserGlnasnilepheleut 137
      459 TTCTCGAAGAGTCATGCATAGATATTAACCAAGCTATGTCTCATTA 508
      137 hrlYaspglyThrValglnleuglYasPheglyllealargValleu 153
      509 CAGCCACTGGGGGTGTAACCTTGAGATCTGGGCTGGCGGTTTTC 558
      154 AsnSerThrValglnleualargThrcysIleglyThrProtyrTyle 170
      559 ACCTCAAAAACCCACAGCTGCACATCTTGTAGTGGTAGCGCTTATCA 608
      170 userProgluilecysgluasnlysproyrasnasnlySeraspillet 187
      609 GTCTCCAGAGCAATACATGAATAATGATACACTTAATCAATTCGATCT 658
      187 rPalaleuIcysvalleuylrgluueucysThrlleuylshsIalphe 203
      659 GGCTCTGGCTGCTACTATATAGATGAGGCTGCATTCACAAAGTCTTTC 708
      204 GluAlaIglySerMetlys....AsnleuValleuylleleSerGI 218

```

```

709 TATGTTGACAAATGATTTACTCTGCTGTAAGAAAGATAGACAGT 758
      ::::::::::::::::::::
218 yserpHeProvaal...SerleuH1STySerTyAspLeuArgSerI 234
      ::::::::::::::::::::
759 TGACACACCACCTCTCTCTCTGATCATTATGAGAACCTCCGACAGT 808
      ::::::::::::::::::::
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
      ::::::::::::::::::::
809 TAGTTAAATATGTCATCAACCCAGATCCAGAGAACGACCAAGCTGAC 858
      ::::::::::::::::::::
251 SerleuGlnLysGlyPheIleAlaLysArgIle 262
      ::::::::::::::::::::
859 TATGTTATGAC.....GTAGCAAGAGGATG 885
      ::::::::::::::::::::

```

seq_name: /cgn2.6/ptodata/1/lna/5B.COMB.seq:US-09-221-237-3

```

seq_documentation_block:
: Sequence 3, Application US/09221237
: Patent No. 6214597
: GENERAL INFORMATION:
: APPLICANT: Acton, Susan
: TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
: FILE REFERENCE: NMI-050
: CURRENT APPLICATION NUMBER: US/09/221,237
: EARLIER FILING DATE: 1998-12-28
: EARLIER APPLICATION NUMBER: 09/163,115
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 3
: LENGTH: 906
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(906)
: US-09-221-237-3

```

alignment_scores:

Quality:	485.50	Length:	262
Ratio:	2.503	Gaps:	6
Percent Similarity:	74.046	Percent Identity:	38.931

alignment_block:

US-09-783-320-4 x US-09-221-237-3 ..

Align seg 1/1 to: US-09-221-237-3 from: 1 to: 906

```

8 GlnLysIleGlyGlySerPheGlyLysAlaIleLeuValLysSerTh 24
      ::::::::::::::::::::
112 AAGAAATTTGGTCGGGACATTTAGTGAAGTTTAGACAGCCTGCTCT 161
      ::::::::::::::::::::
24 GlnAspGlyArgGlnTyValIleLysGlnIleAsnIle...SerArgM 40
      ::::::::::::::::::::
162 CTTCGATGAGTACACAGTCTTAAAAAAAGTCAGATTTTGATTTGA 211
      ::::::::::::::::::::
40 eLSeSerLysGlnArgGlnLysSerArgArgGlnValAlaLysAla 56
      ::::::::::::::::::::
212 TGGATGCCAAAGCACGTCGTGATTCATCAAAAGAAATAGATCTTCT 261
      ::::::::::::::::::::
57 AsnMetLysHisProAsnIleValAlaGlnTyArgGlnSerPheGln 73
      ::::::::::::::::::::
263 CAACACCAACCATCAAAATGTAATAAATATTATGATCATTCATGAGA 311
      ::::::::::::::::::::
73 nGlySerLeuTyTrIleValMetAspTyArgGlnGlyGlyAspLeu... 88
      ::::::::::::::::::::
312 TAAATGACTAAACATAGTTTGGAACTAGACGATGCTGGCAGCTATCCA 361
      ::::::::::::::::::::
89 ....PheLysArgIleAsnAlaGlnLysGlyValLeuPheGlnGlnAsp 103
      ::::::::::::::::::::
363 GAATGATCAACGATTTTAAAGCAAAAGAG...CTAATTCTCTGAAGA 408
      ::::::::::::::::::::

```

```

104 GlnIleLeuAspTrpPheValGlnIleCysLeuAlaLysValHisValH 120
      ::::::::::::::::::::
409 ACTGTTTGAAGATATTGTTGTCAGCTTTCAGTGCATTGGAAACATGACA 458
      ::::::::::::::::::::
120 sAspArgLysIleLeuHisArgAspIleLysSerGlnAsnIlePheLeu 137
      ::::::::::::::::::::
459 TTCGCAAGAGTATGATCATGAGATATAAACCACCTAATGCTCATTA 508
      ::::::::::::::::::::
137 hTrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAlaArgValLeu 153
      ::::::::::::::::::::
509 CAGCCACTGGGGTGTAACCTTGAGATCTTGCGCTGGCCGGTTTTC 558
      ::::::::::::::::::::
154 AsnSerThrValGlnLeuAlaArgThrCysIleGlyThrProTyTrpLe 170
      ::::::::::::::::::::
559 AGCTCAAAAACACACAGCTGCACATTTCTTTAGTGGTACGCCCTTATTA 608
      ::::::::::::::::::::
170 uSerProGlnIleCysGlnAsnLysProTyArgAsnAsnLysSerAspIle 187
      ::::::::::::::::::::
609 GTCTCCAGAGAGATATACATGAATAATGATACAACTTCAAACTGACATCT 658
      ::::::::::::::::::::
187 rPalLeuGlyCysValLeuTyArgLeuCysThrLeuLysHisAlaPhe 203
      ::::::::::::::::::::
659 GGCTCTTGCTGCTCTATATATGAGATGGCTGCATTCACAAAGCTCTTTC 708
      ::::::::::::::::::::
204 GlnAlaGlySerMetLys.....AsnLeuValLeuLysIleIleSerG 218
      ::::::::::::::::::::
709 TATGTTGACAAATGATTTATCTACTCTGCTGTAAGAACATAGAACAGT 758
      ::::::::::::::::::::
218 yserPheProvaal...SerleuH1STySerTyAspLeuArgSerI 234
      ::::::::::::::::::::
759 TGACTACCAACCTCTCTCTCTGATCATTATGAGAGAACCTCCGACAGT 808
      ::::::::::::::::::::
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
      ::::::::::::::::::::
809 TAGTTAAATATGTCATCAACCCAGATCCAGAGAACGACCAAGCTGACC 858
      ::::::::::::::::::::
251 SerleuGlnLysGlyPheIleAlaLysArgIle 262
      ::::::::::::::::::::
859 TATGTTATGAC.....GTAGCAAGAGGATG 885
      ::::::::::::::::::::

```


Wed, May 15 14:14:01 2002

us-09-783-320-4.rn1

THIS PAGE BLANK (USPTO)

1 ATGGAGAGTATGTAGACTACAGAAAGATTGGAGAAAGCTTCATTTGGAAA 50
17 sAlaIleuValIysSerThrGluAspGlyArgInIleValIleIleYSG 34
|||||
51 AGCCATTCTGTTAAATCTACAGAACATGCGACAGTATGTTATCAAGG 100
34 IuIleAsnIleSerArgMetSerSerIysGluArgGluGluSerArgArg 50
|||||
101 AAATTAACTCTCAGATCTCCAGATCTCCAGATCAAGAAAGAAATCAAGAGA 150
51 GluValAlaValIleuAlaAsnMetIysHisProAsnIleValGlnTyrAr 67
|||||
151 GAAGTTCAGATTGGCAACATGAAAGCATCCAAATATTGTCAGATAG 200
67 gGluSerPheGluGluAsnGlySerLeuTyrIleValMetAspTyrCysG 84
|||||
201 AGAATCTTTGAAGAAAATGCTCTCTACATAGTAATGATGACTGTG 250
84 IuGlyGlyAspLeuPheIysArgIleAsnAlaGlnIleYSGIleValLeuPhe 100
|||||
251 AGGAGGCGATCTGTTAAGCGAATAAATGCTCAGAAAGCGCTTTGTTT 300
101 GlnGluAspGlnIleLeuAspTrpPheValGlnIleCysIleuAlaLeuIy 117
|||||
301 CAAGAGCATCAGATTTGGACTGTTTGTACAGATATGTTGGCCCTGAA 350
117 sHisValHisAspArgIysIleLeuHisArgAspIleIysSerGlnAsnI 134
|||||
351 ACATGTCATGATAGAAAATCTTCATCGACATTAATCTCAGAACCA 400
134 IepheLeuThrIysAspGlyThrValGlnLeuGlyAspPheGlyIleAla 150
|||||
401 TATTTTAACTAAAGATGACAGACAGTACAACTTGAGATTTTGCAATGCT 450
151 ArgValLeuAsnSerThrValGluLeuAlaArgThrCysIleGlyThrPr 167
|||||
451 AGAGTCTTAAATAGTACTGTAGAGCTGCTCGAACTTGATGAGGAGCCCC 500
167 oTyTrpLeuSerProGluIleCysGluAsnIysProTyrAsnAsnIys 184
|||||
501 ATACTATTGTCACTGAAATCTGTGTAACAAACCTTACAAATTAATAAA 550
184 eAspIleTrpAlaLeuGlyCysValLeuIyrcIuLeuCysThrLeuIys 200
|||||
551 GTGACATTTGGCTGGCTGGCTGTCTTATGACCTGTGTAACCTTAAA 600
201 HisAlaPheGluAlaGlySerMetIysAsnLeuValLeuIysIleIleSe 217
|||||
601 CATCTTTTGAAGCTGGCAGTATGAAAAACCTGTACTGACATTAATATC 650
217 rGlySerPheProValSerLeuHisTyrSerTyrAspLeuArgSerI 234
|||||
651 TGGATCTTTTCCACCTGTCTTGTGATTAATTCCTATGATCCCGCAGAT 700
234 euValSerGlnLeuPheIysArgAsnProArgAspArgProSerValAsn 250
|||||
701 TGGTGTCTAGTATTTTAAAGAAATCTTAGAGGATAGACATCAGCAAC 750
251 SerIleLeuGlnIysGlyPheIleAlaIysArgIleGluIysPheLeuSe 267
|||||
751 TCCATATTTGAGAAAGGTTTATAGCCAAACGATTGAAAGTTCTCTC 800
267 rProGlnLeuIleAlaGluIuPheCysLeuIysThrPheSerIysPheG 284
|||||
801 TCCACACCTTATTCAGAGAAATTTGTCTAAATAACATTTTGCAGATTGG 850
284 IysSerGlnProIleProAlaIysArgProAlaSerGlyGlnAsnSerIle 300
|||||
851 GATCAGACCTATACAGCTAAAGAACAGCTTCAAGACAAACACGAT 900
301 SerValMetProAlaGlnIysIleThrIysProAlaIalaIysTyrGlyI 317
|||||
901 TCTGTATGCTGCTCAGAAATTAACAAAGCTGCGCTAAATATAGAAAT 950

317 eProLeuAlaTyrIlyIysTyrGlyAspIysIysLeuHisGluIysIysP 334
|||||
951 ACCTTAGCATATTAAGAAATATGAGATTAATAATTCACGAAAGAAAC 1000
334 rLeuGlnIysHisIysGlnAlaHisGlnThrProGluIysArgValAsn 350
|||||
1001 CACTGCAAAACATTAACAGCGCCCATCAACTCCAGAGAAAGAGTGAAT 1050
351 ThrGlyGluGluArgArgIysIleSerGlnIuAlaIlaArgIysArgAr 367
|||||
1051 ACTGGAGAAAGAAAGAGAAATATCTGACGACACAGCAAGAAAGAAAG 1100
367 gLeuGluPheIleGluIysGluIysGlnIysAspGlnIleIleSerI 384
|||||
1101 GCTGAAATTTATTAAGAAAGAAAGAAAGAAAGATCAGATTATAGTT 1150
384 euMetIysAlaGluIuMetIysArgGlnGluIysGluArgLeuGluArg 400
|||||
1151 TAATGAAGCGCTGACAAATGAAAAAGCAAGAAAGAAAGGTTGAAAGA 1200
401 IleAsnArgAlaArgGluGlnGlnIyTrpArgAsnValLeuSerAlaGlyI 417
|||||
1201 ATTAATATAGGCGCCAGGAAACAAAGATGAGAAATGCTAAAGTGTGTC 1250
417 ySerGlyGluValIysAlaProPheLeuGlySerGlyIyThrIleAlaP 434
|||||
1251 AAGTGTGAAGTAAAGCGCTCTTTCTGCGCAGTGAAGGACATTAATTTT 1300
434 rSerSerPheSerSerArgGlyGlnTyrGluHisTyrHisAlaIlePhe 450
|||||
1301 CATCACTTTTCTTCTCTGAGGACAGTATGAACATTACCATGCCATTTT 1350
451 AspGlnMetGlnGluIuArgAlaGluAspAsnGluAlaIysTrpIysAr 467
|||||
1351 GACCAATATGACAGCAACAGACAGACAGATATGAGCTTAATGAAAG 1400
467 gGluIleIyrcIyArgGlyLeuProGluArgGlnIysGlnLeuAlaV 484
|||||
1401 AGAAATATATGTCGAGCTTCGAGAAAGCAAAAGGCGACTACTG 1450
484 aIgluArgAlaIysGlnValGluIuPheLeuGlnIyArgGluAla 500
|||||
1451 TAGAAAGAGCTTAACAAAGTAAAGAGTTCTGCAAGCAAAAGGGAAGCT 1500
501 MetGlnAsnIysAlaArgAlaGluIuGlnIysMetValTyrLeuAlaArgLe 517
|||||
1501 ATGCAGAAATAAAGCTCGACCCGAGACATATGTTTATCTGCAAGACT 1550
517 uArgGlnIleArgLeuGlnAsnPheAsnGluArgGlnGlnIleIysAla 534
|||||
1551 GAGGCAATTAAGACTTACGAATTTCAATGAGCGCAACAGATTAAGCA 1600
534 yIleuArgGlyGluIysIysGluAlaAsnHisSerGluGlyGlnGluIy 550
|||||
1601 AACTTCTGCTGAAAGAAAGAAAGCTAATCATTCGAAAGCAAGAAAGA 1650
551 SerGluGluAlaAspMetArgArgIysIysIleGluSerLeuIysAlaH 567
|||||
1651 AGGAGAGAGGCTGACATGAGCGCAAAAAAATCGAATCTAGAAAGGCCCA 1700
567 sAlaAsnAlaArgAlaIaIaValLeuIysGluIuLeuGluArgIysArgL 584
|||||
1701 TGCAAATATGACGTGCTGCTACTAAAGAACCACTAGAACGAAAGAGAA 1750
584 yGluAlaIyrcIyArgGluArgGluIysIysValITrPGluGluHisIleValAla 600
|||||
1751 AGGAGGCTTATGAGAGAAAAAAGTGGGAAAGACATTTGCTGCT 1800
601 IysGlyValIysSerSerAspValSerProProLeuGlyGlnHisGlyIuTh 617
|||||
1801 AAGAGGTTTAAGGTTCTGATGTTTCTCCACCTTTGGGACAGCATGAAC 1850

617 TGIYGLYSERPROSERLYSGINLMETARGSERVALILESERVALTHRS 634
1851 AGGTGGCTCTCATCAAGCAAGATGATCTGTATTCTGTAACCTT 1900
634 ERALALEUYSGLUVALGLYVALASPSERLEUTHRASPTHRARGLU 650
1901 CAGCTTTGAAAGAAAGTTGGCGTGGACAGTAGTTAACTGATCAGCCGGGAA 1950
651 THRSEGLUMETGLNLYSTRASNANALILESERLYSARGGL 667
1951 ACTTCAGAGAGATGCAAAAAGCAACAAATGCTATTCAAGTAAGCGAGA 2000
667 ULLEUARGHARGLEUASNGULASNULEUYSALAGNULASPGULYSG 684
2001 AATGCTGACAGATTAAATCAAAATCTTAAAGCTCAAGAAAGATGAAAAAG 2050
684 LYMETGLASNULEUSERASPTHRPHGLULLEASNVALHISGLUASPLA 700
2051 GAATGACAGAAATCTCTGATGACTTGTGAGATTAAGTTTCAATGACATGCC 2100
701 LYSGLUHLISGLULYSGLYSSERVALSERSETPARGLYLYSTPGL 717
2101 AAAGAGCATGAAAAAGAAAAATCAGTTTCATCTGATCGCAAGAGTGGGA 2150
717 UALAGLYGLINLEUVALILEPROLEUASPGULUETHRLEUASPHTS 734
2151 GGCAGAGAGTCACACTGTGATTCCTCGATGAGTTAACACTGATACAT 2200
734 ERPHESERTHTRGILUARGHISTRHVALGLYGLUVALILEUYSLEUGLY 750
2201 CCTCTCTACACTGAAGACATACAGTGGGAGAAATTAAATAGCT 2250
751 PROASNGLYSERPROARGARGALATRPGLYLYSERPROTHRASPSEVA 767
2251 CCTAATGATCTCCAAGAGAGCCGGGGGAAAAAGTCCGACAGATCTCTG 2300
767 LLEUYSILLEUGLYGLUVALAGLULEUGNLEUGNLEUGNLEUGNLEUG 784
2301 TCTAAAGATCTTGGAGAAAGCTGAACTCAACTCAGACAGAACTATTAG 2350
784 LUASNTHRITLLEARGSERGLUILESERPROGLUGLYLTYSTYLYS 800
2351 AAAATACAACTATTAAGATGAGATTTCTCCGAAAGGGAAAAAGTACAAA 2400
801 PROLEULEITHRGLYGLULYSVALGLNYSILLESERHISGLULLEAS 817
2401 CCCTTAATTACTGAGAAAAAAAGTACAAATGATTTCTCAGTAAGATAAA 2450
817 NPROSERALALEVALASPSERPROVALGLUETHRYSERPROGLUPHES 834
2451 CCCATCAGCTATTGTTGATTCTCTGTGAGACAAAAAGTCCGAGTTCA 2500
834 ERGLUALASERPROGLIMETSERLEUYSLEUGLUGLYASNLEUGLUGLU 850
2501 GTGAGGCATCTCCACAGATGTCATGAAACTGGAAGAAATTTAGAGAA 2550
851 PROASPSERLEUGLUTHRGULILEUGNGLUPROSERGLYTHRASNYL 867
2551 CCTGATCATTTGGAAACAGAAATTTACAAAGAGCCCAAGTGGAAACAA 2600
867 SASPGULSERLEUPROCYSSTRILETHRSPVALTRPILLESERGLUGLU 884
2601 AGATGAGAGCTTCCATGCACTATTACTGATGTGTGATTTAGTGGAAA 2650
884 YSGULTHRILYSGLUTHRGINSERVALASPARGLIETHRILEGLNULASN 900
2651 AAGAAAAAAGAAAGAACTCAGTCGCGAGATAGGATCAACCATTCAGAAAT 2700
901 GLUVALISERGLUASPGLYVALISERSTRHVALASPGULNLEUSERASPIL 917
2701 GAAGTTCTGAGAGTGAAGTCTGAGTACTGTGAGCAAACTTAGGACAT 2750
917 EHISLILEGLUPROGLYTHRASNASPSERGLINHSERLYSCYASPAVAL 934

2751 TCATRTAGAGCTGGAGAACAAATGATTTCTCAGCACTCTAAATGATGATAG 2800
934 SPYLSERVALGLNPROGLUPROPHENHISLYSVALHISERGLU 950
2801 ATAACTCTGTCAACCGAAACCAATTTTCCATAGAGTGGTTGATTTCTGAA 2850
951 HISLEUASNLEUVALPROGLINVALGLINSERVALGLNLYSERPROGLUGL 967
2851 CACTTGAACCTTAGACCTCCAGATTCATTCAGTTCATCTTCCAGGAA 2900
967 USERHEALAPHEARGSERHISSTRHISLEUPROPOLYASNYL 984
2901 ATCCCTTCATTTGATGATCTCCTGACATTCACCAACAAAAATAAACAA 2950
984 YASNASERLEUENLILEGLYUENSERTHRGLYUENPHEASPALAASN 1000
2951 AGAATTCCTGCTGATGCTGACCTTCAACTGCTGTTGATGCAAAACAC 3000
1001 PROLYSMETLEUARGTHRHCYSERLEUPROASPLEUSERLYSLEUPHEAR 1017
3001 CCAAGATGTTAAGACATGTTTCACTTCAGATCTCTCAAAAGCTGTTGAG 3050
1017 GTHRLEUMETASPPVALPROTHRVALGLYASPPVALARGLINASPSANLEUG 1034
3051 AACCTTATGATGATGTTCCACCGTAGAGAGATGTCGCAAGACAACTTG 3100
1034 LULLEASPGULILEUYSASPGULASNILEUYSGLUGLYPROSERASPSE 1050
3101 AATATGATGAAATTAAGATGAAAAACATTAAGAAAGACCTTCTGATTC 3150
1051 GLUASPILEVALPHEGLUGLUTHRASPSTRHSPLEUINLEUGNGL 1067
3151 GAAGACATTTGTTTGAAGAACTGACACAGATTTTACAAAGCTGCGAGG 3200
1067 ASERMETGLUGNLEUENLEUARGGLUGNINPROGLYGLUGLUTHYSERGLN 1084
3201 CTCGATGGAACAGTTTACAGGAAACACTGGTGGAAATACAGTGAAG 3250
1084 LUGLUGLUSERVALILEUYSASNSEASPPVALGLUPROTHRVALAASNGLY 1100
3251 AAGAAAGCTCAGCTTGAAGAACAGTGAATGTGGACCAACTGCAAAAGGG 3300
1101 THRASPVALALASPGULASPPASPNPROSERSEGLUSERALALEUAS 1117
3301 ACAGATGTGCGAGATGAAAGATGACAAATCCAGTATGTAAAGTCCGTA 3350
1117 NGULGLUTRPHISERASPSASPPVALGLULLEALASERGLUCYSG 1134
3351 CGAAAGATGGCACTCAGATTAACAGTGAATGCTGTAAGTAAGTGTG 3400
1134 LUCYASPSERVALPHEASNHISLEUGLUGLULEUARGLEUENLEUGLU 1150
3401 AATGCGATGATGCTTTAAACATTTAGAGCACTGAGACTTCATCTGAG 3450
1151 GLNGLUMETGLYPEGULYSPHERPEGULUVALITYGLULYSILEYSAI 1167
3451 CAGGAAATGGCTTTGAAAAATTTCTTGAGGTTTGAAGAAATAAAGGC 3500
1167 ALIHISGLUASPPULASPPULASNILEGLULILECYSERLYSILEYALG 1184
3501 TATTTCAGAGATGAGATGAAATATTAATTAATTTGTTCCAAAAATAGTTC 3550
1184 INASNILEUGLYASNGLINHISGLNHSILEUTYVALALYSILEUENHIS 1200
3551 AAAATATTTTGGAAATGAACATCACCATCTTATGAGCAATTTCTTCAT 3600
1201 LEUVALMETALASPPGLYALATYRGINGLUNASPSANSPGLU 1214
3601 TTAGTCATGCGCAGATGAGAGCTTCCAAAGAAAGTAATGATGA 3642
seq_name: /net/abs06/sids1/3c/gdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AA511582

seq_documentation_block:
 ID AAS11582 standard; cDNA: 5426 BP.
 AC AAS11582;
 DT 24-OCT-2001 (first entry)
 XX Human cDNA encoding novel human protein, NHP #26.
 DE Human cDNA encoding novel human protein, NHP; ss; breast cancer; prostate cancer;
 KW Immunogen; antibody; gene therapy; antisense.
 OS Homo sapiens.
 PN W0200161016-A2.
 XX 23-AUG-2001.
 PD 15-FEB-2001; 2001MO-US05356.
 XX 18-FEB-2000; 2000US-0183582.
 FR 22-FEB-2000; 2000US-0184014.
 XX (LEXI-) LEXICON GENETICS INC.
 PA Walke DW, Hu Y, Nepomichy B, Turner CA, Zambrowicz B;
 PI WPI; 2001-502793/55.
 DR Isolated nucleic acids encoding novel human proteins useful for the
 XX treatment of disease and as probes for testing and detection -
 PT Disclosure: Page 68-69; 69pp; English.
 PS The invention relates to novel human proteins (NHP) and the nucleic
 CC acids encoding them. The nucleic acids encode mammalian transporter
 CC proteins and are useful for the treatment (e.g. by gene therapy or
 CC antisense technology) of any of a wide variety of symptoms associated
 CC with biological disorders (e.g. breast and prostate cancer) or imbalances
 CC and as probes for the identification, selection and validation of novel
 CC molecular targets for drug discovery. The proteins may be used to raise
 CC anti-NHP antibodies. The present sequence encodes an NHP of the
 CC invention.
 XX
 SQ Sequence 5426 BP; 1852 A; 927 C; 1163 G; 1484 T; 0 other;

alignment_scores:
 Quality: 6243.00 Length: 1214
 Ratio: 5.143 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-783-320-4 x AAS11582 ..
 Align seg 1/1 to: AAS11582 from: 1 to: 5426

```

1 MetGluGlySerValArgLeuGlnLysIleGlyGluGlySerPheGly 17
  |||||
534 ATGGAGAAATGATGTTAGACTACAGAAATTGGAGAGAGTTCATTGGAAA 563
  |||||
17 salalileuValLysSerThrGluAspGlyArgGlnTyValIleLysG 34
  |||||
584 ACCCATCTCTGTTAAATCTACAGAAAGATGGACAGACTATGTATTCAGG 633
  |||||
34 luileAsnIleSerArgMetSerSerLysGluArgGluGluSerArgArg 50
  |||||
634 AAATTACATCTCCAAAGATGTCAGTAAAGAAAGAAAGATCAAGCAGA 683
  |||||
51 GluValAlaValIleuAlaAsnMetLysHisProAsnIleValGlnTyArg 67
  |||||
684 GAAGTTGCAGTATGCGCAAAACATGAGACATCAATATATGTCCAGTATAG 733
  |||||

```

```

67 gGluSerPheGluGlnLysAsnGlySerLeuTyriIleValMetAspTyrcysG 84
  |||||
734 AGAATTCATTTGAGAAAATGCGCTCTCTACATAGTAAGATTCCTGTG 783
  |||||
84 luGlyGlyAspLeuPheLysArgIleAsnAlaGlnLysGlyValLeuPhe 100
  |||||
784 AGCGAGCGGATCTGTTAAACCAATTAATGCTCAGAAAGCCCTTTGTTT 833
  |||||
101 GlnGluAspGlnIleLeuAspTrpPheValGlnIleCysLeuAlaLeuLys 117
  |||||
834 CAAGAGATCAGATTTGGACCTGTTGTACAGATATGTGTTGCCCTGAA 883
  |||||
117 SHISValHisAspArgLysIleLeuHisArgAspIleLysSerGlnAsnI 134
  |||||
884 ACATGTACATGATAGAAAAATTTCTCATCGACACATTAATCTCAGAAC 933
  |||||
134 IepheLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAla 150
  |||||
934 TATTTTAACTAAAGATGACACAGTACAACTTGAGATTTTGGAATGCT 983
  |||||
151 ArgValLeuAsnSerThrValGluLeuAlaArgThrCysIleGlyThrPr 167
  |||||
984 AGAGTTCTTAATAGTACTGACGCTGCGCTGCAACTGTGATGGGACCC 1033
  |||||
167 OTYrTYrLeuSerProGlnIleCysGluAsnLysProTyraAsnLysS 184
  |||||
1034 ATACTACTGTGCACCTGAAATCTGTGAAAAACAACCTTACAAATTAATA 1083
  |||||
184 eAspIleTrpAlaLeuGlyCysValLeuTyrgluLeuGlySerThrLeuLys 200
  |||||
1084 GTGACATTTGGCGCTCGGGGTGTCTCTTATGAGCGTGTACTACTTAA 1133
  |||||
201 HisAlaPheGluAlaGlySerMetLysAsnLeuValLeuLysIleIleS 217
  |||||
1134 CATCTTTTGAAGCTGGCAATGAAAAACCTGTACTGAAATATATATC 1183
  |||||
217 rGlySerPheProValSerLeuHisTyrseryrAspLeuArgSerL 234
  |||||
1184 TGGATCTTTTCACCTGTGTTCATTAATCTCTGATCCAGATT 1233
  |||||
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
  |||||
1234 TGGTGTCTCAGTATTTAAAGAAATCTAGGATAGACCACTCACTCAAC 1283
  |||||
251 SerIleLeuGlnLysGlyPheIleAlaLysArgIleGluLysPheLeuS 267
  |||||
1284 TCCATATTTGGAGAAAGCTTTTATAGCCAAACCATTTGAAAGCTTCTC 1333
  |||||
267 rProGlnLeuIleAlaGlnLysPheCysLeuLysThrPheSerLysPheG 284
  |||||
1334 TCCACACTTATTCAGAAAGAAATTTGTCTAAAAACATTTTCGAGATTG 1383
  |||||
284 LysSerGlnProIleProAlaLysArgProAlaSerGlyGlnAsnSerIle 300
  |||||
1384 GATCAGAGCCTATACCGCTAAAGAACAGCTTCAGGACAAACATCGATT 1433
  |||||
301 SerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyrglyT 317
  |||||
1434 TCTGTTTTCCTGCTCGAANAATTTACAAAGCCTGCCGCTAAATATGCA 1483
  |||||
317 eProLeuAlaTyriLysLysTyrglyAspLysLysLeuHisGluLysLysP 334
  |||||
1484 ACCTTTACATATTAAGAAATATGAGATTAATAAAATTAACGGAAGAAC 1533
  |||||
334 rGluLeuGlnLysHisLysGlnAlaHisGlnThrProGluLysArgValAsn 350
  |||||
1534 CACTGCAAAAAACATTAACAGCCCATCAAACTCCAGAGAAAGAGATGAA 1583
  |||||
351 ThrGlyGluGluArgArgLysIleSerGluGluAlaAlaArgLysArgAr 367
  |||||
1584 ACTGAGAGAGAAAGAGAAATATCTGAAGAAAGCAAGCAAAAGAAAG 1633
  |||||
367 gLeuGluPheIleGlnLysGlnLysLysGlnLysAspGlnIleIleSerL 384
  |||||

```

```
|||||1634 GCTGCAATTTTAAAAAGAAAAAGAAAGAAAGATCATATTATAGTT 1683
|||||384 eumellysAlaGluGlnMetLysArgGlnGlySerGluArgLeuGluArg 400
|||||1684 TAATGAAGGCTGACAAATGAAAGCAAGAAAGAAAGAAAGTTGGAAAGA 1733
|||||401 IleasnArgAlaArgGlnGlnGlyTyrPargAsnValIleuSerAlaGly 417
|||||1734 ATAAATAGGGCCAGGAACAAAGATGGAGAAATGCTAGTCTGCTGGTGG 1783
417 ySerGlyGluValLysAlaProPheLeuGlySerGlyThrIleAlaP 434
|||||1784 AACGTGTGAAGTAAAGGCTCTTTCTGTGGCAAGTGAGAGGACTATAGCTC 1833
434 roSerSerPheSerSerArgGlyGlnTyrGlnIleuIlePhe 450
|||||1834 CATCATCTTTTCTTCTCGAGAGACAGTATGAACATACCATGCAATTTT 1883
451 AspGlnMetGlnGlnGlnArgAlaGlnLysPasnGluAlaLysTrpLysAr 467
|||||1884 GACCAATGCGCCACAAAGACAGCAAGATATGAACTAAATGCAAAAG 1933
467 gGluIleTyrGlyArgGlyLeuProGluArgGlnLysGlyIleuAlaV 484
|||||1934 AGAAATATATGTCGAGGCTTCCAGAAAGCAAAAGGCAAGCTAGCTG 1983
484 alGluArgAlaLysGlnValGlnGluPheLeuGlnArgLysArgGluAla 500
|||||1984 TAGAAAGAGCTAAACAAAGTAGAAGATTCTCGACAGCAAAAGGGAACCT 2033
501 MetGlnAsnLysAlaArgAlaGlnGlnGlnIleuMetValTyrLeuAlaArgLe 517
|||||2034 ATGCAGATTAAGCTCGAGCCGAAGACATATGCTTATCTGCAAGACT 2083
517 uArgGlnIleArgLeuGlnAsnPheAsnGlnuArgGlnGlnIleLysAlaL 534
|||||2084 GAGGCAATTAAGACTACAGAAATTCATATGAGCCCAACAGATTAAAGCCA 2133
534 yLysLeuArgGlyLysLysGlnLysGluAlaAsnHisSerGlnGlyGlnGly 550
|||||2134 AACCTTCGTGTGAAAAAGAAAGAAAGCTATATCTGAGAGCAAGAGCA 2183
551 SerGlnGluAlaAspMetLarGArgLysLysIleGlnuSerLeuLysAlaH 567
|||||2184 AGTGAAGAGGCTGACATGAGCGCAAAAGAAATGCAATCAGTGAAGCCCA 2233
567 sAlaAsnAlaArgAlaAlaValIleuLysGlnGlnLeuGluArgLysArgL 584
|||||2234 TGCAAATGACGCTGCTCTCTACTTAAAGAAACAATGAAAGCAAGAGAA 2283
584 ySGluAlaArgGlnuArgGlnLysLysValTyrGlnGlnIleuIleuValAla 600
|||||2284 AGGAGCTTATGAGAGAGAAAAAGAGTGGGAAAGCAATTTGGTGGCT 2333
601 LysGlyValLysSerSerAspValSerProProLeuGlnGlnHisGluTh 617
|||||2334 AAGGAGTTAAGAGTTCTGATGTTCTCACCTTTGGGACAGCAATGAAAC 2383
617 rGlyLysSerProSerLysGlnGlnMetArgSerValIleSerValThrS 634
|||||2384 AGGTGGCTCTCCATCAACCAACAGATGAGATCTGTAATTTCTGTACTT 2433
634 eArgIleuLysGlnuValGlyValAspSerSerLeuThrAspThrArgL 650
|||||2434 CAGCTTTGAAAGAGTTGGCGTGGACAGTAACTTAAGTAAACCCGGAA 2483
651 ThrSerGlnuMetGlnIleuThrAsnAsnAlaIleSerSerLysArgG 667
|||||2484 ACTTCAGAAAGATGCAAAAGCAACAAATGCTATTTCAAGTAAAGCGGA 2533
667 uIleLeuArgArgLeuAsnGlnLysLeuLysAlaGlnGlnLysPglLysG 684
|||||

2534 AATACTGCGCAGATTAAATGAAATCTTAAAGCTCAAGAGATGAAAG 2583
684 LyMetGlnAsnLeuSerAspThrPheGlnIleAsnValHisGluAspAla 700
|||||2584 GAATGCGAATCTCTGTGACTTTTGAGTAATGTTCAATGAAGATGCC 2633
701 LysGlnHisGlnLysGlnLysSerValSerSerAspArgLysLysTrpG 717
|||||2634 AAGAGCATTGAAAGAAAGAAATCACTTTCATCTGATGTCGCAAGAAAGTGGCA 2683
717 uAlaGlyGlnGlnLeuValIleProLeuAspGlnLeuThrLeuAspThrS 734
|||||2684 GCGAGAGGTCAACTGTATTCCTGTGATGATGATGATACATACATATCAT 2733
734 ePheSerThrThrGluArgHisThrValGlyGluValIleLysLeuGly 750
|||||2734 CCTTCTGTACACTGAAAGACATACAGTGGGAGGAAATTAATTAATAGT 2783
751 ProAsnGlySerProArgArgAlaTrpGlyLysSerProThrAspSerVa 767
|||||2784 CCTAATGATCTCCAAAGAGAGCTGGGGAAGAAAGTCCGACAGATTTCTGT 2833
767 IleuLysIleLeuGlnGluAlaGlnLeuGlnLeuGlnIleThrGluLeuG 784
|||||784 LAsnThrThrIleArgSerGluIleSerProGlnGlyGluLysLys 800
|||||2884 AAATACACATTTAGAGTGAAGTATTCCTCCGAAGGGGAAAGTACAAA 2933
801 ProLeuIleThrGlyLysLysValGlnGlnIleSerHisGlnIleAs 817
|||||2934 CCTTAATTTACTGGAGAAAAAGATACAAATGATTTCCATGAAATATA 2983
817 nProSerAlaIleValAspSerProValGlnThrLysSerProGluPheS 834
|||||2984 CCCATCAGCTATGTTGTATGTTCTCTGTGACCAAAAGTCCCGATTCA 3033
834 eArgIleLysSerProGlnMetSerLeuLysLeuGlnGlnLysLeuGlnG 850
|||||3034 GTGAGGCATCTCCACAGATGATGATGAACTGACAGGAAATTTAGAGAA 3083
851 ProAspAspLeuGlnuArgLysIleLeuGlnGlnuArgProSerGlyThrAsn 867
|||||3084 CCTGATGATTTGGAACAGAAATTTCTACAAAGCAAGAGTGAACAAACA 3133
867 sAspGlnuSerLeuProCysThrIleThrAspValTrpIleSerGlnGlu 884
|||||3134 AGATGAGAGCTTGCATCAGTCACTATTTACTGATGATGTGATTAAGAGAA 3183
884 ySGluThrArgGlnuArgGlnuSerAlaAspArgLleThrIleGlnGluAsn 900
|||||3184 AAGAAACAAAGAAAGCACTAGTGGCAGATGAGATCACCATTGCGAAAT 3233
901 GluValSerGluAspGlyValSerSerThrValAspGlnLeuSerAsp 917
|||||3234 GAAGTTCTGAAGATGAGATCGAGTCTGTGACCACTTAACTTAACTGAT 3283
917 eHisIleGlnuProGlyThrAsnAspSerGlnHisSerLysCysAspVal 934
|||||3284 TCAATTAAGAGCTGGAAACCAATGATTTCCAGACCTTAATGATGATAG 3333
934 sPrySerValGlnProGluProPhePheHisValValIleHisSerGlu 950
|||||3334 ATAACTGTGCAACCGCAACCAATTTTCCATTAAGGTGTTCAATTCGAA 3383
951 HisLeuAsnLeuValProGlnValGlnSerValGlnCysSerProGlu 967
|||||3384 CACTTAACCTTAATGCTCAATGATTCATATGATGATGATGATGATGAT 3433
967 uSerPheAlaPheArgSerHisSerHisLeuProProLysAsnLysAsn 984
|||||3434 ATCCTTGCATTTGATGATCTGATGATGATGATGATGATGATGATGAT 3483
```

984 ysaasSerleuLeuIleGlyLeuSerThrGlyLeuPheaspAlaasn 1000
 |||||
 3484 AGAATTCCTTGCTGATGTGACCTTCAACTGCTGCTTGTGATGCAACAC 3533
 |||||
 1001 ProLysMetLeuArgThrCysSerLeuProaspLeuSerLysLeuPhe 1017
 |||||
 3534 CCAAAAGATTAAAGCAGCATGTTCACTCCAGATCTCTCAAAAGCTGTTCAG 3583
 |||||
 1017 gThreLeuLeaspValProThrValGlyAspValArgLeuaspLeu 1034
 |||||
 3584 AACCTTATGATGATGTTCCACCGTAGAGATGTTCTGCAAGACATCTTG 3633
 |||||
 1034 LuIleaspGluIleLysaspGluasnIleLysGluGlyProSeraspSer 1050
 |||||
 3634 AAATAGATGAATTTAAAGATGAAAACATTAAGAAGACCTTCTGATTC 3683
 |||||
 1051 GluaspIleValPheGluGluThrAspThrAspLeuGluIleuGln 1067
 |||||
 3684 GAAGACATTTGTTTGAAGAACTGACACGATTTACAGAGCTGAGGC 3733
 |||||
 1067 aserMetGluLeuLeuLeuArgGluGlnProGlyGluGlySerGlu 1084
 |||||
 3734 CTCGATGGAACAGTTACTTAGGAAACAACCTGTGAAGAAATACAGTAA 3783
 |||||
 1084 LuGluGluSerValLeuLysasnSeraspValGluProThrAlaasnGly 1100
 |||||
 3784 AAGAAGAGTCAGTCTTGAAAGAACAGTGTGAGCCCACTGCAAAATGGG 3833
 |||||
 1101 ThraspValAlaaspGluaspAspAsnProSerSerGluSerAlaLeu 1117
 |||||
 3834 ACAGATGTGCGCAGATGAGATGACATCCAGTAGTGAAAGTGCCTGAA 3883
 |||||
 1117 nGluGluThrPheHisSeraspAsnSeraspGlyGluIleAlaSerGluGly 1134
 |||||
 3884 CGAAGAAATGGCAGCTCAGATTAACAGTGAAGTGAATTCGTAAGTAACTG 3933
 |||||
 1134 LuCysaspSerValPheasnHisLeuGluGluLeuArgLeuHisLeuGlu 1150
 |||||
 3934 ATTCGATAGTGTCTTTTACCATTTAGAGAACGAGACATTCATCTGAG 3983
 |||||
 1151 GlnGluMetGlyPheGluLysPheGluValIlyrGluLysIleLysAl 1167
 |||||
 3984 CAGGAAATGGCTTGTGAATAATCTTTCAGGTTTATAGAAAATAAGGC 4033
 |||||
 1167 alLeuHisGluaspGluaspGluasnIleGluIleCysSerLysIleVal 1184
 |||||
 4034 TATTCATGAGATGAGATGAAATATTTGAAATTTGTTCAAAATACTTC 4083
 |||||
 1184 InAsnIleLeuGluasnGluHisLeuIleuGluIleValAlaLysIleLeuHis 1200
 |||||
 4084 AAAATATTTTGGGAAATGACATCGCATCTTTATGCCAAGATCTTCAT 4133
 |||||
 1201 LeuValMetAlaaspGlyAlaIleTyrgluGluaspAsnaspGlu 1214
 |||||
 4134 TTAGTCATGCGCAGATGAGACCTTACCAAGACATATATATGAA 4175
 |||||

seq_name: /net/abss06/STDS1/gcgcdata/hold-geneseq/gene-seq-emb1/NA2001A.DAT-AA158367

seq_documentation_block:

ID AA158367 standard; cDNA; 5448 BP.

AA158367;

22-OCT-2001 (first entry)

Human polynucleotide SEQ ID NO 570.

Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
 peripheral nervous system; neuropathy; central nervous system; CNS;
 Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 amyotrophic lateral sclerosis; Shy-drager Syndrome; chemotactic;
 chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KW Leukaemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AU, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR MPI: 2001-442253/47.
 DR P-PSDB: NAM39211.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Claim 1; SEQ ID NO 570; 10078bp; English.

XX
 CC 'The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (NAM38642-NAM42213) with noctropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

Sequence 5448 BP; 1862 A; 932 C; 1168 G; 1486 T; 0 other;

alignment_scores:

Quality: 6237 00 Length: 1214
 Ratio: 5.142 Gaps: 0
 Percent Similarity: 99.918 Percent Identity: 99.918

alignment_block:

US-09-783-320-4 x AA158367 ..

Align seg 1/1 to: AA158367 from: 1 to: 5448

1 MetGluLysTyrValArgLeuGlnIleGlyGluGlySerPheGlyLys 17
 |||||
 534 ATGGAAGAAGTATGTAGACTACAGAAAGATGGAGAGTTCATTGGAAA 583
 |||||
 17 SAAlleuValLysSerThrGluaspGlyArgGlnTyrValIleLysG 34
 |||||
 584 AGCCATTCCTGTAAATCTACAGAAAGATGGCAGACAGTATGTATCAAG 633
 |||||
 34 LuIleasnIleSerArgMetSerSerLysGluaspGluaspArgArg 50
 |||||

651 ThrSerGluMetGlnLysThrAsnAlaIleSerSerLysArgI 667
 |||||
 2484 ACTTGAGAAAGAGTCAAAAGACCAACATGCTATTTCAGTAAGCGAGA 2533
 667 uIleuArgArgLeuAsnGluAsnLeuLysAlaGlnGluAspGluLysG 684
 |||||
 2534 AATACCTGCTAGATTAATGAAAAATCTTAAGCTCAAGAAAGATGAAAAAG 2583
 684 LysGlnAsnLeuSerAspThrPheGluIleAsnValHisGluAspAla 700
 |||||
 2584 GAAAGCAGATCTCTGATACCTTTTGAGTAATGTTCAATGAAGAGGCC 2633
 701 LysGlnHisGluLysGluLysSerValSerSerAspArgLysArgI 717
 |||||
 2634 AAGAGCATGAAAAAAGAAATCAGTTTCATCTGATCCAGAGAGTGGGA 2683
 717 uAlaGlyGlnLeuValIleProLeuAspGluLeuThrLeuAspThrS 734
 |||||
 2684 GCGAGAGGTCAACTGTGATTCCTGAGATGAGTTAACTGATACAT 2733
 734 e-PheSerThrThrGluArgHisThrValGlyGluValIleLysLeuGly 750
 |||||
 2734 CCTTCTCAACACTGAAAGACATACAGTGGAGAGAAAGTTAAATAGGT 2783
 751 ProAsnGlySerProArgArgAlaTrpGlyLysSerProThrAspSerVa 767
 |||||
 2784 CCTAATGATCTCCAAAGAACCCGGGGGAAAGTCCGACAGATCTGT 2833
 767 IleuLysIleLeuGlyGluAlaGluLeuGlnLeuGlnThrGluLeuLeuG 784
 |||||
 2834 TCTAAAGACTTGGAGAAAGCTGAACCTCACTCAGACAGAACTATTAG 2883
 784 LysAsnThrThrIleArgSerGluIleSerProGluGlyLysArgLys 800
 |||||
 2884 AAAATACAACTTAGAAGTGAATTTCTCCGAAAGGGAAGAAATACAA 2933
 801 ProLeuIleThrGlyGluLysValGlnCysIleSerHisGluIleAs 817
 |||||
 2934 CCTTATATTACTGAGAAAAAAGTACATGATTTTCACATGAAATAAA 2983
 817 n-ProSerAlaIleValAspSerProValGluThrLysSerProGluPheS 834
 |||||
 2984 CCCATCAGCTATTGTGATTTCTCCGTGTGAGACAAAAAGTCCGAGTTCA 3033
 834 e-ArgLysAlaSerProGlnMetSerLeuLysLeuGlnGluLysAsnLeuGlnGlu 850
 |||||
 3034 GTGAGCAGATCTCCACAGATGTCATTGAAACGTAAGAAATTTAGAAAGAA 3083
 851 ProAspAspLeuGluThrGluIleLeuGlnGluProSerGlyThrAsnLys 867
 |||||
 3084 CCTGATGATTTGAAACAGAAATTTCTACAAAGGCCAAAGTGACAAACAA 3133
 867 sAspGluSerLeuProCysThrIleThrAspValTrpIleSerGlnGluL 884
 |||||
 3134 AGATGAGACTTGCCATGCACATTAAGTGTGATGATTAAGTAGAGAAA 3183
 884 ySGluThrLysGluThrGlnSerIleAspArgIleThrIleGlnGluAsn 900
 |||||
 3184 AAGAAACAAAGAAACTCAGTCGCGAGATAGATCACCATTCCAGAAAT 3233
 901 GluValSerGluAspGlyValSerSerThrValAspGlnLeuSerAspI 917
 |||||
 3234 GAAGTTCTGAAGATGAGGTCTCGAGTACTGTGACCAACTTACTGTCAT 3283
 917 eHisIleGluProGlyThrAsnAspSerGlnHisSerLysCysAspAla 934
 |||||
 3284 TCATATAGAGCCTGGAAACCAATGATTCACACACTCTAAAGTCAATAG 3333
 934 sPlySerValGlnProGluProPhePheHisLysValValHisSerGlu 950
 |||||
 3334 ATTAAGTCTGTCCAAACCGAACCATTTTCCATTAAGGTGCTCATTTGAA 3383
 951 HisLeuAsnLeuValProGlnValGlnSerValGlnCysSerProGluG 967

|||||
 3384 CACTTGAACCTTAGTCCCTCAAGTTCAATCAGTTCAGTTCACCGAAGA 3433
 967 userPheAlaPheArgSerHisSerHisLeuProProLysAsnLysAsnL 984
 |||||
 3434 ATCTTTGCAATTCAGATCTCAGTCGATTTACCAACAAAAATTAAGACA 3483
 984 yAsnSerLeuLeuIleGlyLeuSerThrGlyLeuPheAspAlaAsnAsn 1000
 |||||
 3484 AGATTCTCTTGATGTGAGCTTCAACCTGAGAGATGCTGTCACAAACAC 3533
 1001 ProLysMetLeuArgThrCysSerLeuProAspLeuSerLysLeuPheAr 1017
 |||||
 3534 CCAAGATGTTAAGACATGTTTCACTTCCAGATCTCTCAAAAGCTTTGAG 3583
 1017 gThrLeuMetAspValProThrValGlyAspValArgGlnAspAsnLeuG 1034
 |||||
 3584 AACCTTATGAGTGTCCACCGTAGAGATGCTGTCACAAAGCAATCTTG 3633
 1034 LysLeuAspGluIleLysAspGluAsnIleLysGlnGlyProSerAspSer 1050
 |||||
 3634 AAATAGATGAATTAAGATGAAGAAACATTAAGAAAGAGACCTTGATTC 3683
 1051 GluAspIleValPheGluGluThrAspThrAspLeuGlnGluLeuGlnAl 1067
 |||||
 3684 GAAGACATTTGTGTTGAAGAACTGACACAGATTTTACAGAGCTGCAGGC 3733
 1067 aSerMetGluGlnLeuLeuArgGluGlnProGlyGluGlyLysSerGluG 1084
 |||||
 3734 CTCGATGAACAGTACTTATGAGAACACCTGAGTGAAGATACAGTGAAG 3783
 1084 LysGluGluSerValIleLysAsnSerAspValGluProThrLysAsnGly 1100
 |||||
 3784 AAGAAGATCAGCTTTGAACAACATGATGTGAGGACCACTCAATAGGG 3833
 1101 ThrAspValAlaAspGluAspAspAsnProSerSerGluSerAlaLeuAs 1117
 |||||
 3834 ACAGATGTGGAGATGAGATGATGATGATGATGATGATGATGATGATG 3883
 1117 nGluGluThrHisSerAspAsnSerAspGlyGluIleAlaSerGlyCysG 1134
 |||||
 1134 LysAspSerValPheAsnHisLeuGlnGluLeuArgLeuHisLeuGlu 1150
 |||||
 3934 AATCGATAGTGTCTTAAACCATTTAGAGAACTGAGACTTCATCTGAG 3983
 1151 GlnGluMetGlyPheGluLysPhePheGluValIleGlyLysIleLysAl 1167
 |||||
 3984 CAGCAAAATGGCGTTGAAAAATTTCTTGAAGTTATGAGAAAAATTAAGC 4033
 1167 aIleHisGluAspGluAspGluAsnIleGluIleLysSerLysIleValG 1184
 |||||
 4034 TATTCATGAAGATGAAGATGAAGAAATTTGAAATTTGTTCAAAAAATAG 4083
 1184 LysAsnIleLeuGluLysGlnHisGlnHisLeuThrLysLysIleLeuHis 1200
 |||||
 4084 AAAATATTGTTGGAAATGAAGACATCAGCATCTTATGCCAAAGTTTTCAT 4133
 1201 LeuValMetAlaAspGlyAlaIleArgGlnGluAspAspAspGlu 1214
 |||||
 4134 TTAGTCATGGCAGATGGAGCTTCCAAAGAGATTAAGATGAAGA 4175

seq_name: /net/abss06/SDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AA158366
 seq_documentation_block:
 ID AA158366 standard: cDNA: 5532 BP.
 XX
 AC AA158366:
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide seq ID NO 569.


```

1484 ACCCTTACCATAATGAAATATGAGATAAAAAATTACAGAAAGAAC 1533
334 roleuGlnLysHisLysGlnAlaHisGlnThrProGluLysArgValAsn 350
1534 CACTGCAAAAACATTAACAGGCCCATCAAACTCCAGAGAGAGAGATCAAT 1583
351 ThrGlyGluGluArgArgLysIleSerGluGluAlaIleArgLysArgAr 367
1584 ACTGGGGAAGAAAGAGGAAAATATCTGAGGAGACACAGCAAGAAAGAG 1633
367 GLeuGluPheIleGluLysGluLysGlnLysAspGlnIleIleSerL 384
1634 GCTGGAATTTATGAAAAGAAAAGAAAAGAAAGATCAAGATTATTAGTT 1683
384 eumetLysAlaGluGluMetLysArgGlnGluLysGluArgLeuGluArg 400
1684 TATGTAAGGCTGAACAAATGAAAAGGCAAGAAAAGAAAGGTTGGAAAGA 1733
401 IleAsnArgAlaArgGluGlnGlyTrpArgAsnValLeuSerAlaGlyI 417
1734 ATTAATAGGCGCCAGGAAAGAGATGGAAATGTCTTAAGTCTGTGTGG 1783
417 ySerGlyValLysAlaProPheLeuGlySerGlyIleThrIleAlaP 434
1784 AAGTGTAGAGTAAAGGCTCTTTCTGGCGAGTGGAGGACTATAGCTC 1833
434 roSerSerPheSerSerArgGlyGlnTrpGluHisIleThrIleAlaPhe 450
1834 CATCATCTTTTCTCTCGAGCAGATGAACTTACCATTCATTCATTTT 1883
451 AspGlnMetGlnGlnArgAlaGluAspAsnGluAlaLysTrpLysArg 467
1884 GACCAATGCGAGCAACAAAGAGAGCAAGATATGAACTTAATGGAAAAG 1933
467 gGluIleTrpGlyArgGlyLeuProGluArgGlnLysGlyGluLeuAla 484
1934 AGAAATATATGTGTCAGGCTTCCAGAAAAGGCAAAAAGGCGAGCTAG 1983
484 aIGluArgAlaLysGlnValGluGluPheLeuGlnArgLysArgGluAla 500
1984 TAGAAAGAGCTTAACAAAGTAGAAGAGTTCCTGCGAGCAAAAAGGAG 2033
501 MetGlnAsnLysAlaArgAlaGluGlyHisMet..... 511
2034 ATGCAAGATTAAGCTCGAGCGAAGACATATGGAAATCCTGCAAAAAC 2083
511 ..... 511
2084 GGCAGCTATGTATGAGGAGGAGCCAGCTCTTCAAGAGAGGAGCAAA 2133
512 .....ValTrpLeuAlaArgLeuArgGlnIleArgLeu 522
2134 GAACAAAGAGAGAGAGCTTATCTGGCAAGAGCTGAGGCAAAATAGAG 2183
523 GlnAsnProPheAsnGluArgGlnGlnIleLysAlaLysLeuArgGlyLys 539
2184 CAGATTTTCATGAGCGCCCAACAGATTAAACCAACTCTGTGTGAAAA 2233
539 srySerGluAlaAsnHisSerGluGlyGlnGluGlySerGluGluAlaAs 556
2234 GAAAGAGAGTAACTCTGAAAGGACAGAGAGAGGAGAGGCTGTACA 2283
556 eArgArgLysLysIleGluSerLeuLysAlaHisAlaAsnAlaArgAla 572
2284 TGAGCGCGCAAAAAATCGAATCTGAAAGGCCCATGCAAAATGCAAGCG 2333
573 AlaValLeuLysGluGlnLeuGluArgLysArgLysGluAlaTrpArg 589
2334 CCTGTACTTAAAGAACTAGAACTAGAAAGAAAGAGGCTTATGTAGAG 2383
589 gGluLysLysValTrpGluGluHisLeuValAlaLysGlyValLysSer 606
|||||
2384 AGAAAAAAGGTGTGGAGACAGCATTTGGTGCCTAAAGAGCTTAAGAGT 2433
606 eArgValSerProProLeuGlyGlnHisGluTrpGlyLysProSer 622
2434 CTGATGTTTCTCCACCTTTGGGACAGCATGAAACAGGTGGCTCTGCATCA 2483
623 LysGlnGlnMetArgSerValIleSerValThrSerAlaLeuLysGluVal 639
2484 AAGCAACAGATGAGATCTGTTATTCTGTAACTTCAGCTTGTAAAGAGT 2533
639 IGLValAspSerSerLeuThrAspThrArgGluTrpTrpSerGluMetG 656
2534 TGGCTGTGACAGTAGTTTAACGTATATCCCGGAAACTTCAGAAAGAGATC 2583
656 InLysThrAsnAsnAlaIleSerSerLysArgGluIleLeuArgArgLeu 672
2584 AAAAGACCAACAAATGCTATTTCAGTAGCGAAGAAATACTGTGAGATTA 2633
673 AsnGluAsnLeuLysAlaGlnGluAspGluLysGlyMetGlnAsnLeu 689
2634 AATGAAATCTTAAGCTCAAGAGATGAAAAAGGAAACAGAAATCTCTC 2683
689 rAspThrPheGluIleAsnValHisGluAspAlaLysGluHisGluLys 706
706 LysSerValSerSerAspArgLysLysTrpGluAlaGlyMetGlnLeu 722
2734 AAAATAGCTTTCATCTGATCGCAAGAGTGGAGGAGCAGAGCTCACTT 2783
723 ValIleProLeuAspGluLeuThrLeuAspThrSerPheSerThrThrG 739
2784 GTGATCTCTCGATGAGTAGTAACACTAGATACATCCTTCTTCAACATGA 2833
739 uArgHisThrValGlyGluValIleLysLeuGlyProAsnGlySerPro 756
2834 AAGCATACAGCTGGAGAACTTATTAATTAAGTCTTAATGCAATCTCCA 2883
756 rArgAlaIleTrpGlyLysSerProThrAspSerValLeuLysIleLeuG 772
2884 GAAGAGCTCGGGGAAAGTCCGACAGATTCGTGTTTAAAGTAACTGTGGA 2933
773 GluAlaGluLeuGlnLeuGlnThrGluLeuLeuGluAsnThrThrIleAr 789
2934 GAACCTGACTACAACTTCAGACAGAACTATTAAGAAATTAACAATTAAG 2983
789 gSerGluIleSerProGluGlyLysLysTrpLysProLeuIleThrGly 806
2984 AAGTGAGATTTCTCCGAAAGGGGAAAGTAAACCCCTTAATTAAGTGGAG 3033
806 LysLysLysValGlnCysIleSerHisGluIleAsnProSerAlaIleVal 822
3034 AAAAAAAGTACATGATGATTTCACATGAATTAACCCATCACTATGTT 3083
823 AspSerProValGluTrpThrLysSerProGluPheSerGluAlaSerPro 839
3084 GATTCCTCTGTGAGCAAAAAAGTCCCGAGTTCAAGAGACATCTCCACA 3133
839 nMetSerLeuLysLeuGlnGlyAsnLeuGluTrpAspAspLeuGluTr 856
3134 GATGTCAATGAACTGGAGAGAAATTTAGAGAACTGATGATTGGAAA 3183
856 hrGluIleLeuGlnGluProSerGlyThrAsnLysAspGluSerLeuPro 872
3184 CAGAAATTTCTAACAGAGCCCAAGTGGAACAAACAAATGAGGCTTGGCA 3233
873 CysThrIleThrAspValTrpIleSerGluGluLysGluThrLysGluTh 889
3234 TGCACTATTAAGTATGTGTGATTAAGTGAAGAAAAAGAAAGAAAGAAC 3283
889 rGlnSerAlaAspArgIleThrIleGlnGluAsnGluValIleSerGluAsp 906
3284 TCACTCGCAGATAGATGCATCCATTCAGGAAAAATGAAGTTCTGAAGATG 3333

```

906 lvalserSerThrValAspGlnLeuSerAspIleHisIleGluProGly 922
|||||
3334 GAGCTCGAGTACTGTGACCACTTACATTCATATAGAGCCGGA 3383
923 ThrAsnSerSerGlnHisSerIleCysAspValAspIleSerValGlnP 939
|||||
3384 ACCAATGATCTTCAGACACTTAATGTGATGATAGTAACTGTGCAAC 3433
939 cglupropPheHisValValHisSerGlnHisLeuAsnLeuValP 956
|||||
3434 GGAACCAATTTTCATAGGTGCTTCATTCGAACACTGAACTTACTTC 3483
956 roGlnValGlnSerValGlnCysSerProGlnGluSerPheAlaPheArg 972
|||||
3484 CTCAAGTTCATCATGTTGATGATGTCACGAAAGAAATCTTTCATTTG 3533
973 SerHisSerHisLeuProProlCysAsnIleValAsnSerLeuLeuI 989
|||||
3534 TCTCAGTCGATTTACCCCAAAATAAATAAACAAGAAATCTTGTCTGAT 3583
989 eGlyLeuSerThrGlyLeuPheAspAlaAsnAsnProIleLeuValGT 1006
|||||
3584 TGGACTTTCACAGCTGTCTGTTCATGCAAAACCAAGAGATGTAAAG 3633
1006 hrCysSerLeuProAspLeuSerIleValPheArgThrLeuMetAspVal 1022
|||||
3634 CATGTCACCTCCAGATCTCTCAAAAGCTGTCGAACCCCTTATGATGT 3683
1023 ProThrValGlyAspValAlaArgIleAspAsnLeuGluIleAspGluIle 1039
|||||
3684 CCCACCGTAGAGATGTTGCTCAAGCAAACTTGAATGAGAAATTA 3733
1039 SASpGlnAsnIleLeuGlnGlyProSerAspSerGlnAspIleValPheG 1056
|||||
3734 AGATGAACAATTAAGAGAGACCTTCTGATCTGAAGACATGTGCTTGG 3783
1056 lgluThrAspThrAspLeuGlnGluLeuGlnIleAsnSerGlnGlnLeu 1072
|||||
3784 AAGAACTGACACAGATTTACAAAGAGCTGCAGCTCGATGGAACGTTA 3833
1073 LeuArgGlnGlnProGlnGluGluThrSerGlnGluGlnGlnSerValLe 1089
|||||
3834 CTTAGGAAACACCTGTGTGAGAAATACAGTGAAGAAAGACGACATCTT 3883
1089 ulysAsnSerAspValGluProThrAlaAsnGlyThrAspValAlaAspG 1106
|||||
3884 GAAGAACAGTGTGTGAGCCAACTGCAAAATGGACAGATGTGCGAGATG 3933
1106 lAspAspAsnProSerSerGlnSerAlaLeuAsnGlnGluThrPheIle 1122
|||||
3934 AAGATGACAAATCCAGAGTGAAGTGCCTGAAACGAAAGATGCACTCA 3983
1123 AspAsnSerAspGlyGluIleAlaSerGlnCysGlnCysAspSerValP 1139
|||||
3984 GATTAACAGTGTGTAATAATGCTAGTCAATGTAATGCAATGATGTCTT 4033
1139 eAsnHisLeuGlnGluLeuArgLeuHisLeuGlnGlnIleMetGlyPheG 1156
|||||
4034 TAACCAATTAAGAGAACTGAGACTTCATCTGAGACAGGAATGGCTTGG 4083
1156 lIuysPhePheGlnValIleGlyIleValIleValAlaIleHisGlnAspGln 1172
|||||
4084 AAAATTTCTTGGAGTTTATGAGAAATAAAGGCAATTCATGAGATGAA 4133
1173 AspGlnAsnIleGluIleCysSerIleValGlnAsnIleLeuGlnIle 1189
|||||
4134 GATCAAAATTTGAAATTTGTTCAAAAATAGTTCAAAATATTTTGGGAAA 4183
1189 nGlnHisGlnHisLeuThrAlaIleValIleLeuHisLeuValMetAlaAsp 1206
|||||
4184 TGAACATCAGCATCTTATATGCCAAGATTTCTTCAATTTAGTCAATGCGAGATG 4233

1206 lvalATyrGlnGlnIleAspAsnAspGln 1214
|||||
4234 GAGCTTACCAAGAGATATATGATGAA 4259
seq_name: /net/abs06/STDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AA160152
seq_documentation_block:
ID AA160152 standard; cDNA; 5514 BP.
XX AA160152;
DT 22-OCT-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 4141.
DE
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukemia; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200153312-A1.
PN
XX 26-JUL-2001.
PD
XX 26-DEC-2000; 2000WO-US34263.
PF
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
PR
XX
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
PI
XX WPI: 2001-442253/47.
DR P-PSDB: AAM40996.
DR
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PT
XX Claim 1; SEQ ID NO 4141; 10078bp; English.
PS
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AA442213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localized neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 5514 BP; 1500 A; 1191 C; 943 G; 1880 T; 0 other;

Wed May 15 14:14:00 2002

us-09-783-320-4.rng

Page 12

alignment scores: Quality: 6097.00 Length: 1247
 Ratio: 5.047 Gaps: 6
Percent similarity: 96.872 Percent identity: 96.552

alignment block:
us-09-783-320-4 x AA160152/rev ..

Align seg 1/1 to reverse of: AA160152 from: 1 to: 5514

```
1 MetGluTyrValArgLeuGlnLysIleGluGluSerPheGly 17
|||||
4981 ATGGAAGATGTTAGACTACAGAAAGTTGGAAAGCTTGGAA 4932
17 sAlaIleuValIysSerThrGluAspGlyArgGlnTyrValIle 34
|||||
4931 AGCCATTCTGTAAATCTACAGAAAGTGGCAACATGTTATCAGG 4882
34 IuIleAsnIleSerArgMetSerSerIysGluArgGluSerArg 50
|||||
4881 AAATTACATCTCAAGATGTCAGTAAAGAAAGAAAGATCAAGGAG 4832
51 GluValAlaValLeuAlaAsnMetLysHisProAsnIleValGlnTyr 67
|||||
4831 GAAGTTGAGATATGGCAACATGACATCCAAATTTGTCAGATAG 4782
67 GGuSerPheGluGluAsnGlySerLeuTyrIleValMetAspTyrCys 84
|||||
4781 AGAATCATTTGAAAGAAATGGCTCTCTACATAGTAATGATTAAGT 4732
84 LuGluGluAspLeuPheLysArgIleAsnAlaGlnLysGluValIlePhe 100
|||||
4731 AGGAGGGGATCTGTTAAGCAATAATGCTCAAGAACGCTTTCTT 4682
101 GlnGluAspGlnIleLeuAspTrpPheValGlnIleCysLeuAlaLeu 117
|||||
4681 CAAGAGATCAATTTTGACCTGGTGTACAGATATGTTGGCCGAA 4632
117 sHisValHisAspArgLysIleLeuHisArgAspIleLysSerGlnAsn 134
|||||
4631 ACATGTACATGATGAAATAATTTCTACGACATTAATCTCAGAAC 4582
134 LePheLeuThrLysAspGlyThrValGlnLeuGluAspPheGlyIleAla 150
|||||
4581 TATTTTAACTAAAGATGAAACAGTATACTTGAGATTTTGGAAATGCT 4532
151 ArgValIleAsnSerThrValGluLeuAlaArgThrCysIleGlyThr 167
|||||
4531 AGAGTTCTTAATAGTACTGTAGAGCTGGCTGCACTTGCAATAGGAGCC 4482
167 cTyTrpLeuSerProGluIleCysGluAsnLysProTyrAsnAsnLys 184
|||||
4481 ATACTACTTGCTGCACTGAAATCTGTGAAACAAACCTTACATATATATA 4432
184 eArgAspIleArgAlaLeuGlnCysValLeuTyrGluLeuCysThrLeuLys 200
|||||
4431 GTGACATTTGGGCTCTGGGGGTGTCTTATGAGCTGTGACCTGATAA 4382
201 HisAlaPheGluAlaGlySerMetLysAsnLeuValLeuLysIleIle 217
|||||
4381 CATGCTTTGGAAGCTGGCAGTATGAAAAACCTGTACTGAAGATAAATTC 4332
217 rGlySerPheProValSerLeuHisTyrSerTyrAspLeuArgSer 234
|||||
4331 TGGATCTTTTCCACCTGTGCTTTGCAATTATCTATGATCTCCGCGAGT 4282
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
|||||
4281 TGGTGCTCAGTATTAAGAAATCTAGAGATAGCATCAGTCAAC 4232
251 SerIleLeuGluLysGlyPheIleAlaLysArgIleGluLysPheLeu 267
|||||
4231 TCCATATTGGAAGATTTTATACCAACGATTAAGAAAGTTTCTCTC 4182
```

```
267 rProGlnLeuIleAlaGlu...PheCysLeuLysThr.PheSerLys 282
|||||
4181 TCCACACTTATTCACGAAAGATTTTGTCTTAAAAACATTTTTCGAG 4132
283 PheGlySerGlnProIle...ProAlaLysArgProAlaSerGlyLys 298
|||||
4131 TTGGATCACAAACCTTATACAGCTTAAAGACAGCTTCAGCAAAA 4082
298 nSerIleSerValMetProAlaGlnLysIleThrLysProAlaAlaLys 315
|||||
4081 CTCGATTTCTGTTATGCTGCTCAGAAATTTACAAAGCCTGCCGTTAAT 4032
315 yArgIleIleProLeuAlaTyrLysLysTyrGluAspLysLysLeuIle 331
|||||
4031 ATGGAATACCTTTTACGATTAAGAAATATGAGATTAATAATTACACGA 3982
332 LysLysProLeuGlnLysHisLysGlnAlaHisGlnThrProGluArg 348
|||||
3981 AAGAAACCCACTGCAAAACATTAACAGCCCACTCAACATCCAGAGAG 3932
348 gValAsnThrGluGluGluArgArgLysIleSerGluGluAlaAlaArg 365
|||||
3931 AGTGAATACCTGGGCAAGAAAGAGAAATATCTGAGAGAAAGCAGAGAA 3882
365 ySarArgLeuGluPheIleGluLysGluLysLysGlnLysAspGlnIle 381
|||||
3881 AGAGAGAGCTGGAATTTATGAAAAAGAAACAAACAAAGATCAGATT 3832
382 IleSerLeuMetLysAlaGlnGlnMetLysArgGlnGluLysGluArg 398
|||||
3831 ATTAGTTTATGAAAGCTGTAACAAATGAAAGCAAGAAAGAAAGGTT 3782
398 uGluArgIleAsnArgAlaArgGlnGlnIleTyrPheArgAsnValIleSer 415
|||||
3781 GGAAGAAATATATGAGCCAGGACAGAGATGAGAAATGCTGCTAAGTG 3732
415 laGlyLysArgGlyGluValLysAlaProPheLeuGlnLysSerGlyThr 431
|||||
3731 CTGGTGGAGAGTGGAAGTAAAGCTCTTCTGAGGAGTGGAGTGGAGCT 3682
432 IleAlaProSerSerPheSerSerArgGlyGlnTyrGlnHisAla 448
|||||
3681 ATACTCTCATCATCTTTTCTTCTCGAGGACAGTATGAACATTAACATGC 3632
448 aIlePheAspGlnMetGlnGlnIleArgAlaGluAspAsnGluAlaLys 465
|||||
3631 CATTTTGGACCAATGCAGCACAAGAGCAGACAGATATGAAGCTTAAT 3582
465 rPlySarGluIleTyrGluArgGlyLeuProGluArgGlnLysGlyGln 481
|||||
3581 GGAAAGAGAAATATATATGCTGAGGTCTTCCAAAGGCAAAAGGCGAG 3532
482 LeuAlaValGluArgAlaLysGlnValGlnGluPheLeuGlnArgLysArg 498
|||||
3531 CTAGCTGAGAAAGCTTAACAAGTGAAGAGTCTCTCAGCGAAAGACG 3482
498 gGluAlaMetGlnAsnLysAlaArgAlaGluGlyHisMet..... 511
|||||
3481 GGAAGCTATGAGAAATTAAGCTCGAGCCGAAGGACATATGGAATCTGC 3432
511 ..... 511
3431 AAAACCTGCGAGCTATATGAGAGCGAGCCGAGCTTCAAGAGAGGG 3382
512 .....ValTyrLeuAlaArgLeuArgGlnIle 520
|||||
3381 AAGCAAGAAACAAAGAGAGAGGTTTATCTGCAAGCTCAGAGCAAAAT 3332
520 eArgLeuGlnAsnPheAsnGluArgGlnGlnIleLysAlaLysLeuArgG 537
|||||
3331 AAGACTACAGATTTCAATGAGCGCCCAACAGATTAAGCCAAACTTCGTG 3282
```

```

537  LysGluLysLysGluAlaSerHisSerGluGluGluGluGluSerGluGlu 553
538  |||||
539  3281 GTGAAAGAAAGAGCTATCTGTAAGGACAAAGAGGAAAGGAGAG 3232
540  |||||
541  554 AlaSerMetArgArg-LysLysIleGluSerLeuLysAlaHisAlaSer 570
542  |||||
543  3231 GCTGACATGAGCGCAAAAAATCGAATCACTGAGGCCCATCAATG 3182
544  |||||
545  570 LaaArgAlaAlaValIleuLysGluGluGluGluGluGluGluGlu 586
546  |||||
547  3181 CACGTGCTGCTGACTTAAGAAACAACAGTGAAGGAAAGAGAGGCT 3132
548  |||||
549  587 TyrGluArgGluLysLysValIleProGluGluHisLeuValAlaLysGlu 603
550  |||||
551  3131 TATAGAGAGAGAAAAAGGTGGAGAGCATTTGGTGGCTAAAGAGCT 3082
552  |||||
553  603 LysSerSerAspValSerProGluGluGluGluHisGluGluGlu 620
554  |||||
555  3081 TAAAGATTCTGATGTTCTCCACCTTGGGACGACATGAACAGGTGGCT 3032
556  |||||
557  620 erProSerLysGluGluMetArgSerValIleSerValIleSerAlaLeu 636
558  |||||
559  3031 CTCCATCAAGCAACAGATGAGATCTGTAATTTCTGTACTTCACCTTGG 2982
560  |||||
561  637 LysGluValGluValAspSerSerLeuThrAspThrArgGluThrSerGlu 653
562  |||||
563  2981 AAAGAGATTGGCGTGGACAGTAACTTAACGATACCCGGGAAACTTCA 2932
564  |||||
565  653 uGluMetGluLysThrAsnAsnAlaIleSerSerLysArgGluIleLeu 670
566  |||||
567  2931 AGAGATGCAAAAGACCAACATGCTATTTCAGTAACGAGAAATACCTTC 2882
568  |||||
569  670 rGArgLeuAsnGluAsnLeuLysAlaGluGluGluGluGluGluGlu 686
570  |||||
571  2881 GTAGATTAAATGAAAATCTTAAGCTCAAGAGATGAAAAAGAGCGAG 2832
572  |||||
573  687 AsnLeuSerAspThrPheGluIleAsnValHisGluAspAlaLysGlu 703
574  |||||
575  2831 AATCTCTGTACTTTGAGATTAATGTTCAATGAAAGATGCCAAAGACA 2782
576  |||||
577  703 seGluLysGluLysSerValSerSerAspArgLysLysTrpGluAlaGlu 720
578  |||||
579  2781 TGAAGAAAGAAATCATGTTCTCATCTGATCCCAAGAGTGGAGCGAG 2732
580  |||||
581  720 LysLeuLeuValIleProLeuAspGluLeuThrLeuAspThrSerPhe 736
582  |||||
583  2731 GTCAACTGTGTGATCTCTGATGAGTGAACACTAGATACATCTCTCTCT 2682
584  |||||
585  737 ThrThrGluArgHisThrValGluValIleLysLeuGluProAsnGlu 753
586  |||||
587  2681 ACAACTGAAAGACATACAGTGGAGAAAGTATTAATTTAGTCTCTATGG 2632
588  |||||
589  753 YserProArgArgAlaTrpGluLysSerProThrAspSerValLeuLys 770
590  |||||
591  2631 ATCTCCAGAGAGAGCTGGGAGAAAGTCCGACAGATTCGTCTTAAGA 2582
592  |||||
593  770 LeuLeuGluGluAlaGluLeuGluGluGluGluGluGluGluGlu 786
594  |||||
595  2581 TACTTGAGAGAGCTGAACTACAACTTCAGACAGAACTATTGAAAAATCA 2532
596  |||||
597  787 ThrIleArgSerGluIleSerProGluGluGluGluGluGluGlu 803
598  |||||
599  2531 ACATATTGAAAGTGAGATTCTCCGAGAGGGAAGTACAAACCTTAAT 2482
600  |||||
601  803 ethGluGluLysLysValGluCysIleSerHisGluIleAsnProSer 820
602  |||||
603  2481 TACTGAGGAAAAAGTACAAATGTATTTCAATGAAATTAACCATCAG 2432
604  |||||
605  820 LaIleValAspSerProValGluThrLysSerProGluPheSerGluAla 836
606  |||||
607  2431 CTATTGTGATTTCTCGTTGAGACAAAGTCCGAGTTCAGTAGGCA 2382
608  |||||
609  837 SerProGluMetSerLeuLysLeuGluGluGluGluGluGluGlu 853
610  |||||
611  2381 TCTCCACAGATGCTTGAATCGAAGCAAAATTAGAGAACTGATGA 2332
612  |||||
613  853 PleuGluThrGluIleLeuGluGluProSerGluThrAsnLysAspLeu 869
614  |||||
615  2331 TTGGAGAACAGAAATCTCAAGAGCCCAAGTGGAAACAAACAAAGATGAG 2282
616  |||||
617  870 SerLeuProCysThrIleThrAspValTrpIleSerGluGluLysGlu 886
618  |||||
619  2281 AGCTTGCATGACATATCTAGTGTGTGTGATTAATGAGAAAAAGAAC 2232
620  |||||
621  886 rLysGluThrGlnSerAlaAspArgIleThrIleGluGluGluGlu 903
622  |||||
623  2231 AAAGCAACCTCAGTCGACAGATGATGATCAGATGAGAAATGAGATT 2182
624  |||||
625  903 erGluAspGluValSerSerThrValAspGluLeuSerAspIleHis 919
626  |||||
627  2181 CTGAAGATGAGAGTCTCAGTACGTGTGACCACTTACATGATCAATTA 2132
628  |||||
629  920 GluProGluThrAsnAspSerGlnHisSerLysCysAspValAspLys 936
630  |||||
631  2131 GAGCCTGGAAACCAATGATTTCTCAGCACCTCTAAATGTGATGATVAG 2082
632  |||||
633  936 rValGluProGluProPhePheHisLysValValHisSerGluHisLeu 953
634  |||||
635  2081 TCTGCAACCGGAAACATTTTCCATTAAGGTGTTCACTCTGACACTTGA 2032
636  |||||
637  953 snLeuValProGluValGlnSerValGlnCysSerProGluLeuSerPhe 969
638  |||||
639  2031 ACTTAGTCCCTCAAGTTCATCAAGTTCAGTTCACACAGAAATCTTT 1982
640  |||||
641  970 AlaPheArgSerHisSerHisLeuProProLysAsnLysAsnLys 986
642  |||||
643  1981 GCATTCATCTCCTCCTCCATTTCCACCAAAAAATTAACCAAGATTC 1932
644  |||||
645  986 rLeuLeuIleGluLysSerThrGluPheAspAlaAsnAspProLys 1003
646  |||||
647  1931 CTTCGATGATGACATTTTACCTGCTGTTGATGCAACCAACCCAGAGA 1882
648  |||||
649  1003 etLeuArgThrCysSerLeuProAspLeuSerLysLeuPheArgThrLeu 1019
650  |||||
651  1881 TGTTAAGGACATGTTCACTTCAGATCTCAAGCTGTTCACAAACCTT 1832
652  |||||
653  1020 MetAspValProThrValGlyAspValArgGlnAspAsnLeuGlu 1036
654  |||||
655  1831 ATGGATGTTCCCAACCGTAGAGATGTTGTCAACACAACTCTGAATAGA 1782
656  |||||
657  1036 pGluIleLysAspGluAsnIleLysGluLysProSerAspSerGluAsp 1053
658  |||||
659  1781 TGAATTTGAAGATGAAACATTAAAGAGACCTTCGATTCGAAAGACA 1732
660  |||||
661  1053 LeValPheGluGluThrAspThrAspLeuGluGluGluGluGluGlu 1069
662  |||||
663  1731 TTGTGTTTGAAGAACTGACACAGATTTACAAAGAGCTCCAGGCTCGATG 1682
664  |||||
665  1070 GluGluLeuLeuArgGluGluProGluGluGluLysTrpSerGluGlu 1086
666  |||||
667  1681 GAACAGTACTTAGGAAACACCTGTAAGATTCAGTGAAGAGAGA 1632
668  |||||
669  1086 uSerValLeuLysAsnSerAspValGluProThrAlaAsnGluThrAsp 1103
670  |||||
671  1631 GTGAGTCTTAAGAACAGATGATGTGGAGCAACTCAATAGGACAGATG 1582
672  |||||
673  1103 aAlaAspGluAspAspAsnProSerSerGluSerAlaLeuAsnGluGlu 1119
674  |||||
675  1581 TGGCAATGAAATGACAAATCCAGCAGAGAAAGTCCCTGAAGAGAA 1532
676  |||||
677  1120 TrpHisSerAspAsnSerAspGluGluIleAlaSerGluCysGluCys 1136
678  |||||
679  1531 TGGCACTCAAGTACAGTGAATGTTGCTATGAAATGGAATGGA 1482
680  |||||
681  1136 pSerValPheAsnHisLeuGluGluLeuArgLeuHisLeuGluGluGlu 1153
682  |||||

```



```

|||||
4381 CATGCTTTTGAAGTGGCAGATGAAAAACCTGTACTGAGATTAATATC 4332
217 rGlySerPheProIValSerLeuHisTyrSerTyrAspLeuArgSerL 234
|||||
4331 TGGATCTTTTCCACCTGTGCTTGGATATATCTATGATCTCCGCACTT 4282
234 euValSerGlnLeuPheLysArgAspProArgAspArgProSerValAsn 250
|||||
4281 TGGGTCTCAAGTATTTAAAGAAATCTCAGGATAGACCATCACTCAAC 4232
251 SerLeuGlnLysGlyPheLeuAlaLysArgLLeuLysPheLeuSe 267
|||||
4231 TCCATATTGGAGAAAGSTTTTATAGCCAAACGATTTGAAAGTTTCTCTC 4182
267 rProGlnLeuIleAlaGlnLys...PheCysLeuLysThr.PheSerLys 282
|||||
4181 TCCCTCAGCTATTTGACGAAAGAAATTTTGTCTTAAAAACATTTTTCGAG 4132
283 PheGlySerGlnProIle...ProAlaLysArgProAlaSerGlyGlnAs 298
|||||
4131 TTTGGATCACAAGCCTTATACCAAGCTTAAAAAGCCAGCTTCAGGACAAAA 4082
298 nSerIleSerValMetProAlaGlnLysLLeuThrLysProAlaAlaLysT 315
|||||
4081 CTCGATTTCTGTATGCTGCTGCTGCAAAATTTACAAAGCCTGCCGCTAAAT 4032
315 yfGlyIleProLeuAlaTyrLysLysTyrGlyLysPlysLysLeuHisGlu 331
|||||
4031 ATGGAAATACCTTAGCATATAGAAATATGAGATTAATAAAAAATTACCGAA 3982
332 LysLysProLeuGlnLysHisLysGlnAlaHisGlnThrProGlnLysArg 348
|||||
3981 AAGAAACCATGCTCAAAACATTAACAGCCATCAAACTCCAAAGAGAGAG 3932
348 yAlaAsnThrGlyGlnLysArgLysLleSerGlnGlnAlaAlaArgL 365
|||||
3931 AGGAAATCTGGCGAAAGAAAGAGAAATATCTGAGAGAGCAGCAAGAA 3882
365 yAspArgLysGlnLysPheLLeuLysGlnLysLysGlnLysAspGlnIle 381
|||||
3881 AGGAAAGCTGGCAATTTATGAAAAAGAAAGAAACAAAAGATCGAGATT 3832
382 IleSerLeuMetLysAlaGlnLysMetLysArgGlnLysGlnLysArgL 398
|||||
3831 ATTAGTTTAAATGAGGCTGAAACAAATGAAAGGCAAGAAAGGAAAGTT 3782
398 uGlnArgLLeuAsnArgAlaArgLysGlnGlyTTPArgAsnValLeuSerA 415
|||||
3781 GGAAGATTAATATAGCGCCAGGAAACAAGGATGGAATGCTGCTAAGTG 3732
415 laGlyLysSerGlyGlnValLysAlaProPheLeuGlySerGlyLysThr 431
|||||
3731 CTGGTGGAAAGTGTGAAGTAAAGGCTCTTCTGGGCAAGTGAGGAGACT 3682
432 IleAlaProSerSerPheSerSerArgLysGlnLysTyrGlnHisAla 448
|||||
3681 ATACCTCCATCATCTTTTCTCTCTCGAGACAGTATGAACATTTCCCATGC 3632
448 aIlePheAspGlnMetGlnGlnLysArgAlaGlnLysAspAsnGlnAlaLysT 465
|||||
3631 CATTTTGCACCAATGACCAACAAGACACAGACATATATGATGAGTAAAT 3582
465 rPlysArgGlnLysTyrGlyArgGlyLeuProGlnLysArgLLeuLysGln 481
|||||
3581 GGAAAAAGAAATATATGCTCGAGGCTTCCAGAAAGGCAAAAAGGCGAG 3532
482 LeuAlaValGlnArgAlaLysGlnValGlnLysPheLeuGlnArgLysArg 498
|||||
3531 CTAGCTGAGAAAGAGCTAAACAAGTAAAGAGATTCCTCAGCAGAAAGC 3482
498 gGlnAlaMetGlnAsnLysAlaArgLysGlnGlyHisMet..... 511
|||||

3481 GGAAGCTATGCAGATTAAGCTCGAGCCGAGAGACATATGGAATCTCTGC 3432
511 ..... 511
3431 AAAACCTGGCAGCTATGTATGAGAGCAGGCCAGCTCTCAAGAGGAGG 3382
512 .....ValTyrLeuAlaArgLeuArgGlnIle 520
3381 AAGCCAAAGAAACAAAGAGAGAGGTTTATCTGGCAAGCTCAGGCAAAAT 3332
520 eArgLeuGlnAsnPheAsnGlnArgGlnGlnLysAlaLysLeuArg 537
3331 AAGACTACGAAATTTCAATGAGCGCCACACAGATTAAGCCAAACTGCTG 3282
537 LysGlnLysGlnAlaAsnHisSerGlnGlyGlnGlnLysSerGlnLys 553
3281 GTGAAAGAAAGAAAGCTAATCATCTTGAAGAGCAACAAAGAGAGTGAAG 3232
554 AlaAspMetArgArg.LysLysLLeuLysSerLeuLysAlaHisAlaAsnA 570
3231 GCTGACATGAGCGCCAAAAAAATCGAATCTCGAAGGCCCAATGCAAAATG 3182
570 laArgAlaAlaValLeuLysGlnLysGlnLysArgLysArgLysGlnAla 586
3181 CACGTGCTGCTGTACTAAAGAAACACTAGAACGAAAGAGAGAGAGCT 3132
587 TyrGlnArgGlnLysLysValTTPGlnGlnHisLeuValAlaLysGlyVa 603
3131 TATGAGAGAGAAAAAAAGGTGGAGAGACATTTGGCTGAAGAGACT 3082
603 LysSerSerAspValSerProProLeuGlnLysHisGlnThrGlyLys 620
3081 TAAAGATTCTGATGTTTCTCCACTTTGGGACAGCATGAAGACAGTGCT 3032
620 ePProSerLysGlnLysMetArgSerValIleSerValThrSerAlaLeu 636
3031 CTCATCAAGACACACAGATGAGATCTGTATTTCTGTAACCTCAGCTTG 2982
637 LysGlnValGlyValAspSerSerLeuThrAspThrArgLysThrSerG 653
2981 AAGAAAGTTGGCTGACGATGTTTAACTGATACCCGGAACCTTCAG 2932
653 uGlnMetGlnLysThrAsnAsnAlaIleSerSerLysArgGlnIleLeuA 670
2931 AGAGATGCAAAAGACCAACAAATGCTATTTCAAGTAAAGCGAAGAAATCTTC 2882
670 rArgLysAsnGlnLysAsnLeuLysAlaGlnLysAspGlnLysGlyMetGln 686
2881 GTAGATTAAATGAATCTTAAAGCTCAAGAGATGAAAAAGAAAGCAG 2832
687 AsnLeuSerAspThrPheGlnLysAsnValHisGlnAspAlaLysGlnI 703
2831 AATCTCTGATACTTTTGAATTAATGTCAATGAAAGATGCAAAAGAGCA 2782
703 sGlnLysGlnLysSerValSerSerAspArgLysTTPGlnAlaLys 720
2781 TGAAGAAAGAAATCAAGTTTCACTGATCGCAAGAAAGGAGGAGCAGAG 2732
720 LysGlnLeuValIleProLeuAspGlnLeuThrLeuAspThrSerPheSer 736
2731 GTCAACTTGATCTCTCTCGAGATGAGTAACTGATACATCACTTCTCT 2682
737 ThrThrGlnArgHisThrValGlyGlnValIleLysLeuGlyProAsnG 753
2681 ACAACTGAAGACATACAGTGGGAGAAATTAATTAATTAAGCTCAATAG 2632
753 ySerProArgArgAlaTArgLysSerProThrAspSerValLeuLysI 770
2631 ATCTCCAGAGAGAGCTCGGGGAGAAAGTCCGACAGATTCGTGTAAAGA 2582
770 leLeuGlyGlnAlaGlnLeuGlnLysThrGlnLeuLeuGlnLysThr 786
2581 TACTTGGAGAGCGAAGTCAACTTCAAGCAAGCACTATTAGAAATATACA 2532

```

787 ThrIleArgSerGluIleSerProGluGlyIleuStyTrpLysProIleuI 803
 2531 ACTATTAGAGTGAAGATTCTCCCAAGGGGAAAAAGTACAAACCCCTTAAT 2482
 803 eThrGlyGlyLysValGlnCysIleSerHisGluIleAsnProSerA 820
 2481 TACTGGAGAAAAAAGTACATGATTTCATGAAATAAACCCTACAG 2432
 820 IalIleValAspSerProValGluTrpLysSerProGluPheSerGluAlA 836
 2431 CTAATGTGATTCTCCTGTTGAGACAAAAGTCCCGAGTTCAAGAGCA 2382
 837 SerProGluMetSerLeuLysLeuGluLysAsnLeuGluLysProAspAs 853
 2381 TCTCCACAGATGTCAATGAAACTGGAGAAATTTAGAGAACCTGATGA 2332
 853 PLeuGluTrpGluIleLeuGlnGluProSerGlyTrpAsnLysAspGlu 869
 2331 TTTGGAAACAGAAATTCTACAGACCAGTGAACAAACAAAGATAGA 2282
 870 SerLeuProCysThrIleThrAspValTrpIleSerGluGluLysGluTrp 886
 2281 AGCTTCCATGCACATTAATGATGCTGTGATTAGTGAAGAAAAAGAAC 2232
 886 rLysGluTrpGlnSerAlaAspArgIleThrIleGlnGluAsnGluValS 903
 2231 AAGGAAACTCAGTCGGCAGATAGATCACCATTCCAGGAAAAATGAGTTT 2182
 903 eGluAspGlyValSerSerThrValAspGlnLeuSerAspIleHisIle 919
 2181 CTAAATATGAGTCTGAGTACTGTGACCAACTAGTGAATCATCATATA 2132
 920 GluProGlyThrAsnAspSerGlnHisSerLysCysAspValAspLysSe 936
 2131 GAGCTGGAGAACATGATTTCTAGACACTCTAAATGTGATGATGAAGTC 2082
 936 rValGlnProGluProPhePheHisLysValValHisSerGluHisLeuA 953
 2081 TGTGCAACCGAACCATTCTTCCATAGAGTGTGATTCATTCAGACATTGA 2032
 953 snLeuValProGlnValGlnSerValGlnCysSerProGluGluSerPhe 969
 2031 ACTTAATCCCTCAAGTCAATCAGTTCAGTGTCCACGAAAGAAATCCCTT 1982
 970 AlaPheArgSerHisSerHisLeuProProLysAsnLysAsnLysAsnSe 986
 1981 GCATTTGATCTCCTCCTCGCATTTACCAACCAAAAAATAAACAAGAAATTC 1932
 986 rLeuLeuIleGlyLeuSerThrGlyLeuPheAspAlaAsnAsnProLysM 1003
 1931 CTTGCTGATTTGACCTTCAACTGCTGTTGATGCAAAACCAACCAAGA 1882
 1003 eLeuValArgThrCysSerLeuProAspLeuSerLysLeuPheArgThrLeu 1019
 1881 TGTTAAGACATGTTCACTTCACATCTCTCAAAAGCTTTTACAGAACCTT 1832
 1020 MetAspValProThrValGlyAspValArgGlnAspAsnLeuGluIleAs 1036
 1831 ATGAGATGTTCCACCGTAGAGATGTGTCAAGACAAATCTTGAAGAATGA 1782
 1036 pGluIleLysAspGluAsnIleLysGluGlyProSerAspSerGluAspI 1053
 1781 TGAATTTGAGATGAAACATTAAAGAGACCTTCTGATTTCTGAACAACA 1732
 1053 LeValPheGluGluTrpAspThrAspLeuGlnGluLeuGlnAlaSerMet 1069
 1731 TTTGTTTGAAGAACTGACACAGATTTCACAAAGAGCTCAGGCTCGATG 1682
 1070 GluGlnLeuLeuArgGluGlnProGlyGluGluTrpSerGluGluGluGlu 1086
 1681 GAACAGTTACTTAGGAGCAACCTGTGTGAAGATACAGTGAAGAAGAAAGA 1632

1086 uSerValLeuLysAsnSerAspValGluProThrAlaAsnGlyThrAspV 1103
 1631 GTCAATCTTTGAAGAACAGTATGTGAGCAACTGCAGAAATGGAGACAGATG 1582
 1103 aAlaAspGluAspAspAsnProSerSerGluSerAlaLeuAsnGluGlu 1119
 1581 TGGCAGATGAAGATGACAAATCCACAGAGTGAAGTCCCTGAAACCAAGAA 1532
 1120 TrpHisSerAspAsnSerAspGlyGluIleAlaSerGluCysGluCysAs 1136
 1531 TGGCAGCTCAGATACAGTATGATGTAATTCCTAGTGAATGTGAATGCGGA 1482
 1136 pSerValPheAsnHisLeuGlnGluLeuArgLeuHisLeuGlnGluGlu 1153
 1481 TAGTGTCTTTTACCATTTTACAGAGTCACTTCATCTGGAGCAGAGAA 1432
 1153 eGlyPheGluLysPhePheGluValTrpGluLysIleLysAlaIleHis 1169
 1431 TGGCCTTTGAAAAATTTCTTGAAGTTTATGAGAAATTAAGGCTATTCAT 1382
 1170 GluAspGluAspGluAsnIleGluIleCysSerLysIleValGlnAsnI 1186
 1381 GAAGATGAGATGAAATATTTGAAATTTGTTCAAAAATGTTCAAAAATAT 1332
 1186 eLeuGlyAsnGluHisGlnHisLeuTrpAlaLysIleLeuHisLeuValM 1203
 1331 TTTGGAAATGAAACATCAGCATCTTATGCAAGATTTTCATTTAGTCA 1282
 1203 eAlaAspGlyAlaTrpGlnGluAspAsnAspGlu 1214
 1281 TGGCAGATGAGAGCCTTACCAAGAAATATGATGATA 1247
 seq_name: /net/abs06/SIDS1/9cdata/hold-geneseq/geneseqn-embL/NA2001A.DAT:AA511559
 seq_documentation_block:
 ID AA511559 standard; cDNA; 3024 BP.
 AC AA511559;
 DT 24-OCT-2001 (first entry)
 XX Human cDNA encoding novel human protein, NHP #3.
 DE Human: novel human protein; NHP; ss; breast cancer; prostate cancer;
 KW Immunogen; antibody; gene therapy; antisense.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..3024 /tag= a
 FT /product= "NHP #3"
 XX
 PN MO200161016-A2.
 PD 23-AUG-2001.
 XX
 PF 15-FEB-2001; 2001WO-US05356.
 XX
 PR 18-FEB-2000; 2000US-0183582.
 PR 22-FEB-2000; 2000US-0184014.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Walke DW, Hu Y, Nepomilchy B, Turner CA, Zambrowicz B;
 XX WPI; 2001-502793/55.
 DR P-PSDB; AA007103.
 XX
 PT Isolated nucleic acids encoding novel human proteins useful for the
 PT treatment of disease and as probes for testing and detection -
 XX
 PS Disclosure; Page 40; 69pp; English.

XX The invention relates to novel human proteins (NHP) and the nucleic acids encoding them. The nucleic acids encode mammalian transporter proteins and are useful for the treatment (e.g. by gene therapy or antisense technology) of any of a wide variety of symptoms associated with biological disorders (e.g. breast and prostate cancer) or imbalances CC and as probes for the identification, selection and validation of novel CC molecular targets for drug discovery. The proteins may be used to raise CC anti-NHP antibodies. The present sequence encodes an NHP of the invention.

50 Sequence 3024 BP; 1115 A; 503 C; 681 G; 725 T; 0 other:

alignment_scores: Length: 1007
 Quality: 5162.00 Gaps: 0
 Ratio: 5.126
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-783-320-4 x AAS11559 ..

Align seg 1/1 to: AAS11559 from: 1 to: 3024

```
208 MetLysAsnLeuValLeuLysIleIleSerGlySerPheProProValSe 224
|||||
1 ATGAAAAAAGCTGGTACTGAGATATATCTGGATCTTTTCACCTGCTGTC 50
|||||
224 rLeuHisTyrSerTyrAspLeuArgSerLeuValSerGlnLeuPheLysA 241
|||||
51 TTTGGATATATCTCATGATCTCCGAGTTGGTGTCTCACTTATTTAAAA 100
|||||
241 rGAsnProArgAspArgProSerValAsnSerIleLeuGlnLysGlyPhe 257
|||||
101 GAAATCCATGAGATAGACCATCACTCAATTCATTTGGAGAAAGGTTT 150
|||||
258 IleAlaLysArgIleGlnLysPheLeuSerProGlnLeuIleAlaGln 274
|||||
151 ATAGCCAAAGCATGAAAGTTCTCTCTCAGCTATTCAGAGAG 200
|||||
274 rPheCysLeuLysThrPheSerLysPheGlySerGlnProIleProAla 291
|||||
201 ATTTTGTCTAAACATTTTCGAAGTTGGATGCACAGCTATACACAGCTA 250
|||||
291 yAsnArgProLysSerGlnAsnSerIleSerValMetProAlaGlnLys 307
|||||
251 AAGACCCAGCTTCAGACAAAACCTGATTTCTGTATGCCCTCAGCAAAA 300
|||||
308 IleThrLysProAlaAlaLysTyrGlyIleProLeuAlaTyrLysLys 324
|||||
301 ATTACAAAGCTGCCCTAAATATGAAATACCTTTAGCATATAAGAAATA 350
|||||
324 rGlyAspLysLysLeuHisGlnLysLysProLeuGlnLysHisLysGln 341
|||||
351 TGGAGATAAAAAATTACAGAAAGAGAGTGAATCTGCAAAACATTAACAG 400
|||||
341 IahISglnThrProGlnLysArgValAsnThrGlyGlnGluArgLys 357
|||||
401 CCCATCAAACTCCAGAGAAAGAGTGAATCTGCAAAACATTAACAG 450
|||||
358 IleSerGlnGlnLysAlaAlaArgLysArgLeuGlnPheIleGlnLys 374
|||||
451 ATATCTGAGGAGACAGCAAGAAAGAGAGCTGGAATTTATTGAAAGAA 500
|||||
374 ULysLysGlnLysAspGlnIleIleSerLeuMetLysAlaGlnIleMet 391
|||||
501 AAAGAAACAAAGAGATCATGATTTATGTTAATGAAGCTGAAACAAATGA 550
|||||
391 ySAsnGlnGlnLysGluArgLeuGlnArgIleAsnArgAlaArgGlnGln 407
|||||
551 AAAGCAAGAAAGAAAGAGTTGGAAGATTAATAGGCCGAGGAAACAA 600
|||||
408 GlyTrpArgAsnValLeuSerAlaGlyLysSerGlnValLysAlaPr 424
```

```
|||||
601 GGATGGAAGAAATGCTAAGTCTGAGAGAGGGAAGTGAAGTAAAGCTCC 650
|||||
424 rPheLeuGlnSerGlyGlyThrIleAlaProSerSerPheSerArgG 441
|||||
651 TTTTCTGGGCGAGTGGAGGAGCTAATGCTCATCATCTTTTCTTCTCCAG 700
|||||
441 yGlnIleGlnLysIleTyrHisAlaIlePheAspGlnMetGlnGlnArg 457
|||||
701 GACAGTATGAACATTTACATGCCATTTTTCACCAATTCAGCAACAGAA 750
|||||
458 AlaGlnAspAsnGlnAlaLysTyrPheArgGlnIleTyrGlyArgGly 474
|||||
751 GCAGAAAGATATGAAAGCTAATGGAAGAGAAATATATGCTGAGGCT 800
|||||
474 rProGlnArgGlnLysGlnLeuAlaValGluArgAlaLysGlnValG 491
|||||
801 TCCAGAAAGGCAAAAAGGCGAGCTAGCTGTAAGAAAGCTAAACAGTAG 850
|||||
491 LglLysPheLeuGlnArgLysArgLysAlaMetGlnAsnLysAlaArgAla 507
|||||
851 AAGAGTCTCTGAGCGAAACGGAGAGCTATGCAAGATTAAGCTCAGGCC 900
|||||
508 GlnGlyHisMetValTyrLeuAlaArgLeuArgGlnIleArgLeuGln 524
|||||
901 GAAGGACATATGCTTTATCTGCAAGACTGAGCAAAATAAGACTACAGAA 950
|||||
524 rPheAsnGlnArgGlnGlnIleLysAlaLysLeuArgGlyLysLysG 541
|||||
951 TTTCAATGAGCGCCCAACAGATTAAGCCAAACTCGGTGGAAGAAAG 1000
|||||
541 LysAlaAsnHisSerGlnGlnGlnGlnLysSerGlnGlnLysAlaPheMetArg 557
|||||
1001 AACCTAATATCTCTAGACAGACAGAGAACTGAAGAGGCTGACATGAGG 1050
|||||
558 ArgLysLysIleGlnSerLeuLysAlaHisAlaAsnAlaArgAlaLys 574
|||||
1051 CGCAAAAATGCAATCATGAGGCCCATCAATGACAGTCTCTGCTG 1100
|||||
574 IleLysGlnGlnLeuGlnLysArgLysGlnLysAlaTyrGlnArgGln 591
|||||
1101 ACTTAAAGAAACAACATGAAAGAGAAAGAGAGAGCTTATGAGAGAGAA 1150
|||||
591 ySLysValTyrPglLysHisLeuValAlaLysGlyValLysSerSerAsp 607
|||||
1151 AAAAAGTGTGGAAAGAGCATTTGTGGCTAAAGAGTTAAGATTTCTGAT 1200
|||||
608 ValSerProProLeuGlnLysGlnIleGlnLysSerProSerLysGln 624
|||||
1201 GTTCTTCCACCTTTGGGACAGCATGAACAGGTGCTCTCATCAATAAGCA 1250
|||||
624 rGlnMetArgSerValIleSerValThrSerAlaLeuLysGlnLysGln 641
|||||
1251 ACAGATGAGATCTGTAATTTCTGTAACCTTCAGCTTGAAGAAAGCTGGCG 1300
|||||
641 AlaAspSerSerLeuThrAspThrArgGluThrSerGlnGlnMetGlnLys 657
|||||
1301 TGGACATAGTATTAACTGATACCCGGGAAACTTCAGAGAGATCAAAAG 1350
|||||
658 ThrAsnAsnAlaIleSerSerLysArgGlnLysLeuArgGlyLeuAsnGln 674
|||||
1351 ACCAACAATGCTATTTCAGTAAGGAGAAATACCTTCACAGATTAATGA 1400
|||||
674 uAsnLeuLysAlaGlnGlnLysAspGlnLysGlnMetGlnAsnLeuSerSPT 691
|||||
1401 AAATCTTAAGCTTCAGAGATGAAGAAAGATTCAGAGATCTCTCTGATA 1450
|||||
691 hrPheGlnIleAsnValHisGlnAspAlaLysGlnHisGlnLysGlnLys 707
|||||
1451 CTTTGGATGAATGAATGATGAGATGCAAGAGAGCTGAGAAAGCAAAA 1500
|||||
708 SerValSerSerAspArgLysLysTyrPglLysAlaGlyLysGlnLeuValIle 724
|||||
```


XX 15-FEB-2001; 2001WO-US05356.
 PF 18-FEB-2000; 2000US-0183582.
 PR 22-FEB-2000; 2000US-0184014.
 XX (LEXI-) LEXICON GENETICS INC.
 PA
 PI Walke DW, Hu Y, Nepomichy B, Turner CA, Zambrowicz B;
 XX WPI: 2001-502793/55.
 DR P-PSDB; AA007101.
 XX
 PT Isolated nucleic acids encoding novel human proteins useful for the
 treatment of disease and as probes for testing and detection -
 PS Claim 1; Page 32-33; 69pp; English.
 XX
 XX The invention relates to novel human proteins (NHP) and the nucleic
 acids encoding them. The nucleic acids encode mammalian transporter
 CC proteins and are useful for the treatment (e.g. by gene therapy or
 CC antisense technology) of any of a wide variety of symptoms associated
 CC with biological disorders (e.g. breast and prostate cancer) or imbalances
 CC and as probes for the identification, selection and validation of novel
 CC molecular targets for drug discovery. The proteins may be used to raise
 CC anti-NHP antibodies. The present sequence encodes an NHP of the
 CC invention.
 CC
 SO Sequence 3108 BP; 1145 A; 520 C; 708 G; 735 T; 0 other;

Alignment_scores:
 Quality: 5138.00 Length: 1035
 Ratio: 5.102 Gaps: 1
 Percent Similarity: 97.295 Percent Identity: 97.295

alignment block:
 US-09-783-320-4 x AAS11557

Align seg 1/1 to: AAS11557 from: 1 to: 3108

208 MetLysAsnLeuValLeuLysIleIleSerGlySerPheProProVal 224
 1 ATGAAAACCTGGTACTGAAATATATCTGAGCTTTCCACCTGTGTC 50
 224 rleuHisLysSerLysPheAspLeuArgSerLeuValSerGlnLeuPheLysA 241
 1 TTTGCTATTATCTCATGATCTCCGAGTTTGCTCTCACTTATTTAAAA 100
 241 rGAsnProArgAspArgProSerValAsnSerIleLeuGluLysGlyPhe 257
 101 GAATATCTGAGGATAGACCATCAGCACTCATATTGGAGAAAGGTTT 150
 258 IleAlaLysArgIleGluLysPheLeuSerProGlnLeuIleAlaGlu 274
 151 ATGAGCCAAACGCAATGAAAAGTTTCTCTCTCAGCTTATTTGAGAGA 200
 274 upheCysLeuLysThrPheSerLysPheGlySerGlnProIleProAla 291
 201 ATTTTGTCTAAAAACATTTTCGAACTTTGGATCAGCATATACACCTA 250
 291 ysArgProAlaSerGlyLysAsnSerIleSerValMetProAlaGlnLys 307
 251 AAGACCCAGCTTCAGCAAAACTGCATTTCTGTTATGCCGCTCAAAAA 300
 308 IleThrLysProAlaLysArgLysGlyLysLeuAlaLysLysLys 324
 301 ATTACAAAGCTCCGCTAAATATGAAATACCTTACATATAGAAATA 350
 324 rGlyAspLysLysLeuHisGluLysLysProLeuGlnLysHisLysGln 341
 351 TGGAGATAAATAATTACAGAAAAGAAACCACTGCAGAAAACATTAACAG 400

341 IahISGlnThrProGlnLysArgValAsnThrGlyGluGluArgArgLys 357
 401 CCCATCAACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 450
 358 IleSerGlnGlnLysAlaLysArgLysArgLysPheIleGluLysG 374
 451 ATATCTGAGAGACGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 500
 374 uLysLysGlnLysAspGlnIleIleSerLeuMetLysAlaGlnMet 391
 501 AAGAGAAACAAAGAGATCAGATTATTTAGTTATGAGGCTGACAAATGA 550
 391 ysArgGlnGluLysGluArgLeuArgIleAsnArgAlaArgGln 407
 551 AAGGCAAG 600
 408 GlyTPArgAsnValLeuSerAlaGlyLysGlyGluValLysAlaP 424
 601 GGATGAGAGAAATGTCTAAGTGTGTGAGAGAGAGAGAGAGAGAGAG 650
 424 oPheLeuGlySerGlyLysThrIleAlaProSerSerPheSerArg 441
 651 TTTTCTGGGAGTGGAGGAGCTATAGCTCATCATCTTTCTCTCGAG 700
 441 LysIleTyGlnHisLysIleAlaIlePheAspGlnMetGlnGlnArg 457
 701 GACAGTATGAAACATTCACATCCATTTTGTGACCAAAATGACAGAAAGA 750
 458 AlaGluAspAsnGluAlaLysTrpLysArgGluIleTyGlyArgGly 474
 751 GCAGAAATATGAAAGCTAAATGAGAAAGAAATATATGTGAGGCTT 800
 474 uProGluArgGlnLysGlyGlnLeuAlaValGluArgAlaLysGlnVal 491
 801 TCCAGAAAGGCAAAAGGCGAGCTAGCTGTAGAAAGGCTAAACAGTAG 850
 491 LysLysPheLeuGlnArgLysArgGluAlaMetGlnAsnLysAlaArg 507
 851 AAGAGTCTCGAGCGAAGAACGGAACGATATGCAATATAAGCTCAGCC 900
 508 GluGlyHisMet..... 511
 901 GAAGGACATATGTGAAATCTGCAAAACCTGGCAGCTATGATGAGGAG 950
 512 YALT 513
 951 GCCCAGCTCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTTT 1000
 513 yLeuAlaArgLeuArgGlnIleArgLeuGlnAsnPheAsnGluArgGln 529
 1001 ATCTGGCAAGACTGAGGCAAAATAGACTACAGAAATTCATAGAGCCAA 1050
 530 GlnIleLysAlaLysLeuArgLysGluLysLysGlnLysAsnHisSer 546
 1051 CAGATTAAAGCCAACTGCTGTGTAAGAAAGAAAGAGCTAATCTTCTGA 1100
 546 uGlyGlnGluLysSerGlnGluAlaAspMetArgArgLysLysIleGlu 563
 1101 AGGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1150
 563 eLeuLysAlaHisAlaAsnAlaArgAlaAlaValLeuLysGlnLeu 579
 1151 CAGTGAAGGCCCTGCAATATGACAGCTGCTGCTGCTGCTGCTGCTG 1200
 580 GluArgLysArgLysGlnLysArgLysGlnLysLysValLysPheGlu 596
 1201 GAACGAAAG 1250
 596 uHisLeuValAlaLysGlyValLysSerSerPheLysProProLeu 613
 1251 GCATTTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1300
 613 LysGlnHisGluThrGlyLysSerProSerLysGlnMetArgSerVal 629

```

1301 GAGAGCATAAACAGTGGCTCTCCATCAAAAGCAACAGATGATGTCGTT 1350
630 lIeSerValThrSerAlaLeuLysGluValGlyAlaSPSerSerLeuth 646
1351 ATTTCGTAACTTCAGCTTTGAAAGACATGCGCTGGACAGTACTTTAAC 1400
646 rAspTrArGluThrSerGluGluMetGlnLysThrAsnAlaIles 663
1401 TGTATCCCGGAACCTTCGAAAGAGATGCAAAAGCCAACTATCTATT 1450
663 eSerLysArGluThrLeuArGluLeuAsnGluAsnLeuLysAlaGln 679
1451 CAAGTAGGAGAAATACTTCCAGATTAATGAAATCTTAAAGCTCAA 1500
680 GluAspGluLysGlyMetGlnAsnLeuSerAspThrPheGluIleAsnVa 696
1501 GAGATGAAAGAGATGAGATCTCTGATCTCTGATCTTGAATGAATGT 1550
696 lHisGluAspAlaLysGluHisGluLysGluLysSerValSerSerAspa 713
1551 TCATGAAGATGCCAAAGAGCATGAAAGAAATCAAGTTTCATCTGATC 1600
713 rGlyLysTrpGluAlaGlyGlnLeuValIleProLeuAspGluLeu 729
1601 GCAGAGAGTGGAGGAGAGAGCTCAACTGTGATCTCTGATGAGCTTA 1650
730 ThrLeuAspThrSerPheSerThrThrGluArGlnIleThrValGlyGluVa 746
1651 ACACGTGATACATCTCTCTCTCACTGAAAGACATACAGTGGAGAGT 1700
746 lIleLysLeuGlyProAsnGlySerProArGAlaIleTrpGlyLysSerP 763
1701 TATTAAATTAGTCCCTCAATGATCTCCAAAGAGCCCTGGGGGAAAGTC 1750
763 rOThrAspSerValLeuLysIleLeuGlyValAlaGluLeuGlnLeuGln 779
1751 CGACAGATTCCTCTTAAAGATACCTTGGAGAAAGCTGAACCTCAACTCAG 1800
780 ThrGluLeuLeuGlnAsnThrThrIleArGSerGluIleSerProGluGlu 796
1801 ACAGACACTTGTAGAAATACACTATGTAGAGATGATTCCTCCGAGAG 1850
796 yGluLysTrpLysProLeuIleThrGlyLysValGlnCysIles 813
1851 GGAAGAGTCAAAACCTTAATTAATGAGAGAAAGAAAGTCAATGATATT 1900
813 eHisGluIleAsnProSerAlaIleValAspSerProValGluThrLys 829
1901 CACATGAATTAACCCATCAGCTATGTGTGATCTCCCTGAGACAAAA 1950
830 SerProGluPheSerGluAlaSerProGlnMetSerLeuLysLeuGluGlu 846
1951 AGTCCGAGATTCAGTAGAGCATCTCCACAGATGATGAACAGTGAAGG 2000
846 yAsnLeuGlnGluProAspAspLeuGluThrGluIleLeuGlnLeuProS 863
2001 AAATTTAGAGAACTGATGATTTGGAACAGAAATCTTACAAAGCCAA 2050
863 eGlyThrAsnLysAspGluSerLeuProCysTrpIleThrAspValTrp 879
2051 GTGGAAACAACAAGATGAGCTGACATCCATCTTACTGATGTGTGG 2100
880 lIleSerGluGluLysGluThrLysGluThrGlnSerAlaAspArGileTh 896
2101 ATTAGTAGAGAAAGAAAGAAAGAAAGAACTCAGTCGGCAGATAGAGTAC 2150
896 rIleGlnGluAsnGluValSerGluAspGlyValSerSerThrValAspG 913
2151 CATTCAGAAATGAAGTTTGAAGATGAGTCCGAGTACTGTGGACC 2200
913 lIleuSerAspIleHisIleGluProGlyThrAsnAspSerGlnHisSer 929
2201 AACCTAGTACATTCATATAGACCTTGAAACCAATGATTCAGACACTCT 2250
930 LysCysAspValAspLysSerValGlnProGluProPhePheHisLysVa 946
2251 AAATGTGTATGATTAAGTCTGTGCAACCGGAACCAATTTTTCATPAAGT 2300
946 lValHisSerGlnHisLeuAsnLeuValProGlnValGlnSerValGlnC 963
2301 GGTTCATTCGTGAACACTTGAACCTTAGTCCCTCAAGTTCAATCAGTTCA 2350
963 ySerProGluGluSerPheAlaPheArGSerHisSerHisLeuProPro 979
2351 GTTCACCAAGAAATCTTTCGATTTTCGATTCACCTCGCATTTACCACCA 2400
980 LysAsnLysAsnLysAsnSerLeuLeuIleGlyLeuSerThrGlyLeuPh 996
2401 AAAAATTAATAACAAGAAATCTTGTGATGAGACTTTCACACTGATGTGT 2450
996 eAspAlaAsnAsnProLysMetLeuArGThrCysSerLeuProAspLeuS 1013
2451 TGATGCAAAACCAACCAAGAGTGAAGAGATGTTCACTTCCAGATCTCT 2500
1013 eTrpLysPheArGThrLeuMetAspValProThrValGlyAspValArg 1029
2501 CAAAGCTGTTCCAGAACCTTATGAGATGTTCCACCTAGAGATGTGCT 2550
1030 GlnAspAsnLeuGluIleAspGluIleLysAspGluAsnIleLysGluGlu 1046
2551 CAAGACATCTTGAATATAGATAAATTAAGATGAAGAAACATTAAGAAAG 2600
1046 yProSerAspSerGluAspIleValPheGluGluThrAspThrAspLeuG 1063
2601 ACCTTCGATTCGTAAGACATGTTGTTGAAGAACTGACACAGATTCAC 2650
1063 lGlnLeuGlnAlaSerMetGlnLeuLeuArGlnGlnProGlyGluGlu 1079
2651 AAGAGCTGCAGGCTCCATGGAACAGTACTTGGGAACAACTGCTGTA 2700
1080 GluTrpSerGluGluGluGluSerValLeuLysAsnSerAspValGluPr 1096
2701 GAATTCAGTGAAGAGAGAGAGTCACTGTGAAACAGATGATGTGGACC 2750
1096 oThrAlaAsnGlyThrAspValAlaAspGluAspAspAsnProSerSerg 1113
2751 AACTGCAAATGGGACAGATGTGGCAGATGAAGATGACAATCCACAGTGTG 2800
1113 lUsErAlaLeuAsnGluGluTrpHisSerAspAsnSerAspGlyGluIle 1129
2801 AAAGTCCCTGAACGAGAAGATGGCACCTCAGATTAACAGTATGTGAAAT 2850
1130 AlaSerGluCysGluCysAspSerValPheAsnHisLeuGluGluLeuAr 1146
2851 GCTAGTGAATGTGAATGCGATAGTCTTTAACCATTTAGAGAACTGAG 2900
1146 gLeuHisLeuGlnGlnGluMetGlyPheGluLysPhePheGluValTrpG 1163
2901 ACTTCATCTGAGCAGAGAAATGGCTTTGAAAAATCTTGTGAGCTTATAG 2950
1163 lUlySIlleLysAlaIleHisGluAspGluAspGluAsnIleGluIleCys 1179
2951 AGAAATTAAGGCTATTCATGAGATGAAGATGAATAATTAAGAAATTTGT 3000
1180 SerLysIleValGlnAsnIleLeuGlyAsnGluHisGlnHisLeuTrpAl 1196
3001 TCAAAATATAGTTCAAATATTTTGGGAAATGAACATCACCACATCTTATGC 3050
1196 aLysIleLeuHisLeuValMetAlaAspGlyAlaTrpGlnGluAspAsnA 1213
3051 CAAGATTCCTCATTTAGTCAATGAGATGAGCTTACCAAGAAAGATTAATG 3100
1213 sPglu 1214
3101 ATGAA 3105

```

```
seq_name: /net/abs06/SIDSL/9c9data/hold_geneseq/geneseqn-emb1/NA2000.DAT:AAA09328
seq_documentation_block:
ID   AAA09328 standard; DNA; 4263 BP.
XX
XX   AAA09328:
XX
XX   10-AUG-2000 (first entry)
XX
XX   Human cancer associated antigen precursor DNA, clone NY-REN-55.
DE
XX
XX   renal cancer; cancer associated antigen precursor; diagnosis;
XX   cytostatic; murine NEK1 protein kinase homologue; ss.
XX
XX   Homo sapiens.
OS
XX
XX   WO200020587-A2.
XX
XX   13-APR-2000.
XX
XX   04-OCT-1999; 99WO-US22873.
XX
XX   05-OCT-1998; 98US-0166300.
XX   05-OCT-1998; 98US-0166350.
XX
XX   (LUDW-) LUDWIG INST CANCER RES.
XX
XX   Ohta Y, Gout I, Tureci O, Sahin U, Pfreundschuh M, Scanlan MJ;
XX   Stockert E, Chen Y, Old LJ, Jager E, Knuth A;
XX   WPI: 2000-303774/26.
XX
XX   Preventing, diagnosing and/or treating disorders associated with
XX   abnormal expression of human cancer associated antigens
XX
XX   Claim 57: Page 100-101; 121pp; English.
XX
XX   AAA09321-45 were isolated by SEREX screening from a renal cancer
XX   cell line 1973/10.4. Homology searching revealed that these clones
XX   correspond to known genes. The present sequence has identity with the
XX   murine NEK1 protein kinase homologue gene. The genes encode cancer
XX   associated antigen precursors.
XX   These gene products are useful in methods for preventing,
XX   diagnosing and/or treating disorders, especially cancers, associated with
XX   abnormal expression of human cancer associated antigens. The method
XX   comprises contacting a sample from a subject with an agent that
XX   specifically binds to the nucleic acid molecule or expression product
XX   (or fragment) complexed with a human leukocyte antigen (HLA) molecule
XX   and determining the interaction between the agent and the nucleic acid
XX   molecule or the expression product as a determination of the disorder.
XX
XX   Sequence 4263 BP; 1364 A; 829 C; 1065 G; 1005 T; 0 other;

alignment_scores:
Quality: 5040.50      Length: 1214
Ratio: 4.496          Gaps: 6
Percent Similarity: 92.339      Percent Identity: 81.796

alignment_block:
US-09-783-320-4 x AAA09328 ..

Align seg 1/1 to: AAA09328 from: 1 to: 4263

1 MetGluLysTyrValArgLeuGlnLysIleGlyLugLysSerPheGlyIly 17
|||||
576 ATGAGAGAGATGAGTGAAGCTGAGAAAGATGAGAGAGTTCATTTGGAAA 625
17 saiaiaieuaiaiaiaiaiaiaiaiaiaiaiaiaiaiaiaiaiaiaia 34
|||||
626 AGCTGTTCTTGTAAATGACAGAGAGATGGCAGCATATATGTCAACAG 675
```

```
34 IuileasniIleSerArgMetSerSerIyGluArgGluGluSerArgArg 50
|||||
676 AAATTAACTCTCAAGATGCTGTGATTAAGAAAGCAAGCAATCAAGAGAGA 725
51 GluValAlaValIleuAlaAsnMetLysHisProAsnIleValGlnTyrAr 67
|||||
726 GAAATTGCTGTATGGCAAAATGAAAGCATGCAAAATATGTCATATATA 775
67 gGluSerPheGluGluAsnGlySerLeuTyrIleValMetAspTyrCysG 84
|||||
776 AGAATCATTTGAAGAAATGCTCTCTCATAGTAATGATTCCTGTG 825
84 IugIyGlyAspLeuPheLysArgIleAsnAlaGlnLysGlyValIlePhe 100
|||||
826 AAGGAGGTGATTTGTTTAAACGAATTAATCTCGAAGAAAGCCCTCTGTT 875
101 GlnGluAspGlnIleLeuAspTyrPheValGlnIleCysLeuAlaLeuIly 117
|||||
876 CAAAGAAAGACAGATTTTGACTGCTGTGTCAGATATGTTGGCTCTGAA 925
117 shiValHisAspArgLysIleLeuHisArgAspIleLysSerGlnAsnI 134
|||||
926 GCATGTACATGATGATAAAATTTCTTCCACGACATTAAGTCACAGAAC 975
134 IepheLeuThrLysAspGlyThrValGlnLeuGlnLysAspPheGlyIleAla 150
|||||
976 TATTTCTAACCAAGATGGACAGCTGACCTTGAGATTTTGGAAATTCCT 1025
151 ArgValIleuAsnSerThrValGlnLeuAlaArgThrCysIleGlyThrPr 167
|||||
1026 CAGATCTTATATAGTACGTGACGTGACGTGACGTGACGTGACGTGAC 1075
167 ofTyrTyrLeuSerProGluIleCysGluAsnLysProTyrAsnAsnLys 184
|||||
1076 ATACACTGTGTACCTGAAATCTGTGAACCAACCCCTTAAACATAAAA 1125
184 eTAspIleThrAlaLeuGlyCysValLeuTyrGluLeuGlnCysThrLeuLys 200
|||||
1126 GTGACATTTGGCTTTGGCTGTGTGCTGTGTGTGTGTGTGTGTGTGT 1175
201 HisAlaPheGlnAlaGlySerMetLysAsnLeuValLeuLysIleIleSe 217
|||||
1176 CAGGCAATTTGAAGCTGAAACATGAAACCTGTACTGTGAAGTAATCTC 1225
217 rGlySerPheProProValSerLeuHisTyrSerTyrAspLeuArgSerI 234
|||||
1226 CGGATCTCTTCTCCAGTGTCTCCACATTAATCTATGATCTCCGAGCT 1275
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
|||||
1276 TGCTGTCTCAGTTATTTAAAGAAATCTCAGGATGATGACATCATGCAAC 1325
251 SerIleLeuGlnLysGlyPheIleAlaLysArgIleGlnLysPheLeuSe 267
|||||
1326 TCCATATGAGAAAGGTTTATGCTAAACGAATCAAAAGTTTCTCTC 1375
267 rProGlnLeuIleAlaGluPheCysLeuLysThrPheSerLysPheG 284
|||||
1376 CCCTCAGCTTATTCGAAAGATTTTGTCTAAACACTTCAAAAGTTTG 1425
284 IySerGlnProIleProAlaLysArgProAlaSerGlyGlnAsnSerIle 300
|||||
1426 GACACAGCCCTCTCCAGGTAAACACACGATCAGACAGCAAGGTCTAGT 1475
301 SerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyrGlyI 317
|||||
1476 TCTTTGTCCTCTCGAAGAAATCAAAAGCTGCTGTAATATCGAGAGT 1525
317 eProLeuAlaTyrLysLysTyrGlyAspLysLysLeuHisGlnLysLysP 334
|||||
1526 GCCCTTAATATTAAGATATGAGATTAAGAAAGTTACTGTAGAAAAAC 1575
334 rLeuGlnLysHisLysGlnAlaHisGlnThrProGlnLysArgValAsn 350
```

```

1176 CACCCCAAAACATMAACAGCCCATCAATTCCTCGAAGAAAAGAAAT 1625
351 ThrGluGluArgLysIleSerGluGluAlaAlaArgLysArg 367
1626 TCTGGAGAGAAAGAGAAAATGTCTGAGAGACGCAAAAAAGAG 1675
367 gLengUphetIleGluLysGluLysGluLysAspGlnIleIleSer 384
1676 GTTGGAAATTTATGAGAAAGAAAAGCAAAAGATCAG...ATTAGGT 1722
384 eumElySAIAGluGluMetLysArgGluGluLysGluArgLeuGluArg 400
1723 TCTGAAAGCTGACACATGACAGCGCAAGACAGACCGTTGGAGAG 1772
401 IleAsnArgAlaArgGluGluGlyTPArgAsnValLeuSerAlaGly 417
1773 ATAAATAGGCGAGGAAACAGAGATGGAGAAATGTTTAAAGGCTGTGG 1822
417 ySerGlyGluValLysAlaProPheLeuGlySerGlyThrIleAlaP 434
1823 AAGCGGTGAAGTAAAGGCTTCCTTTTGGCATTTGAGAGGCGCTCTCTC 1872
434 roSerSerPheSerArgGlyGlnArgIleuHisTyRHisAlaIlePhe 450
1873 CATCACCTGTCTCTCTCGAGCCAGATGACATTACCATTCGCAATTTT 1922
451 AspGlnMetGluGluArgAlaGluAspAsnGluAlaLysTrpLysArg 467
1923 GACCAATGCGACGCGTAAAGAGCAAGATGAATGAAGCAAGATGAAGG 1972
467 gGluIleArgGlyArgGlyLeuProGluArgGlnLysGluLeuAlaV 484
1973 GGGATCTATGCTGATGCTCCAGAAAAGCAAAAAGCACTTATGCTG 2022
484 aGluArgAlaLysGluValGluGluPheLeuGlnArgLysArgGluAla 500
2023 TAGAGAGGCCAACCAAGTGAAGAAATCTTACAGCGTAAACAGCAAGCT 2072
501 MetGlnAsnLysAlaArgAlaGluGlyHisMetValTyRLeuAlaArg 517
2073 ATGCAAGATTAAGCCGAGCGAGACACGCTGTTATTTGGCAAGACT 2122
517 uArgGlnIleArgLeuGlnAsnPheAsnGluArgGlnIleLysAlaL 534
2123 GAGCAATTAAGACTACAAAATTTTAATGAGCCCAACAGATTAAGCCA 2172
534 ySLeuArgGlyLysLysLysGluAlaAsnHisSerGluGluGluGly 550
2173 AACTCGGTGAGATTAAGAAAGCTGATGTACCAAGACANAGAAAGCA 2222
551 SerGluGluAlaAspMetArgArgLysLysIleGluSerLeuLysAlaH 567
2223 ACTGAAGAGACTGACATGAGCGCTCAAAAAGATGAGTCACTTAAGCGCA 2272
567 sAlaAsnAlaArgAlaValAlaValLeuLysGluGlnLeuArgLysArgL 584
2273 AACAAATGCACTGCTCTCTGACTAAAGAAACAGCTGAGCGCAAAAAGAA 2322
584 ySGluAlaTyRArgLysLysLysValTyRLeuGlnHisLeuValAla 600
2323 AGGAAGCTTATGAAGAAAGAAAGATAGGAAGCAACATTTGGTGCGG 2372
601 LysGlyValLysSerSerAspValSerProLeuGluGlnHisGluThr 617
2373 AGG...GTAAGAACTCAGATGTCTCTCTGCTTGGAACTTCTTAAAC 2419
617 rGlyGlySerProSerLysGluGlnMetArgSerValIleSerValThs 634
2420 AGGTGGTCTTCATCAAGACAGCGAGTGAAGCTGTTCATTTCTGTACTT 2469
634 eAlaLeuLysGluValGlyValAspSerSerLeuThrAspTrpArgGlu 650

```

```

2470 CAGCTTTGAAAGAGTGGCCCTGATGGAAGTTTAATGATACCGAGAA 2519
651 ThrSerGluGluMetGlnLysThrAsnAlaIleSerSerLysArgI 667
2520 .....GAAGAAATGGAAGAAAGAGTAAACAGTCTATTTCAAGTAA 2563
667 uIleLeuArgArgLeuAsnGluAsnLeuLysAlaGlnGluAspGluLysG 684
2564 AATCTGCGGAGGCTTAATGAATCTTAATCAAGCTCAAGAGATGAAAG 2613
684 LysMetGlnAsnLeuSerAspThrPheGluIleAsnValHisGluAspAla 700
2614 AAAGAGCATCTACAGTCTCTTGAGACCGCTTGCTGACAAAGATGAG 2663
701 LysGluHisGluLysGluLysSerValSerSerAspArgLysTyRPl 717
2664 AGAGAGATGAGACAGAAATGCAATTCCTCTGATGCAAGAAAGTGGA 2713
717 uAlaGlyGlyGlnLeuValIleProLeuAspGluLeuThrLeuAspThs 734
2714 GATGGAGGTCACGCTGTGATTCCTCGATGCACTGACACTGATACAT 2763
734 eRPhSerThrThrGluArgHisThrValGlyGluValIleLysLeuGly 750
2764 CTTTCTCTGACACCGAAAAACATACTGTGGAGAGTTATTAATTTGAT 2813
751 ProAsnGlySerProArgArgAlaTyRPlLysSerProThrAspSerVa 767
2814 TCTAATGGCTCTCCAGAAAAGCTGGGGGAAAAACCTTACAGATTCGT 2863
767 IleLysIleLeuGluGlyAlaGluLeuGlnLeuGlnArgGluLeuG 784
2864 GCTGAAGATATCTTGGAAGAGTGAATTAACGCTATAGACAGACTAG 2913
784 LysAsnThrThrIleArgSerGluIleSerProGluGlyLysLysTyR 800
2914 AAACACATCTTTTAAAGTAGAGTTTATGCTGAAGAGAGAACTACAA 2963
801 ProLeuIleThrGlyLysLysValGlnCysIleSerHisGluIleAs 817
2964 CCCTTCTTACAGAGAGAGATTCGACGTCATTTCAAAAGAAATTA 3013
817 nProSerAlaIleValAspSerProValGluThrLysSerProGluPhe 834
3014 TCCATTCAGCTACTGTTGATTTCT...ACTGAAGCAAAAGTCCAAAGTTA 3060
834 eGluAlaSerProGlnMetSerLeuLysLeuGluGlyAsnLeuGlu 850
3061 CTGAGGCTGCTCCCAAAATGTCA.....GAAGGAAATGTGGAAGA 3101
851 ProAspAspLeuGluThrGlnIleLeuGlnGluProSerGlyThrAsn 867
3102 CTTGATGATTTGGAAACAGAGTTCACAAAGAGCCAACTAGACACAC 3151
867 sAspGluSerLeuProCysThrIleThrAspValTyRPlIleSerGlu 884
3152 AGATGGAGATTTCCACGCTTCTTAATGATGTGCGACTGAGAGAGAG 3201
884 ySGluThrLysGluThrGlnSerAlaAspArgIleThrIleGlnGluAsn 900
3202 AAGCAGCTAAGAAAGCTGAGTGAAGATTAAGTTGCTGTGAGAGAGT 3251
901 GluValSerGluAspGlyValSerSerThrValAspGlnLeuSerAsp 917
3252 GAAATTTGTGAAGATGAATTCACAGGAAAGCTGAGCAATCCTTAAGA 3301
917 eHisIleGluProGlyThrAsnAspSerGlnHisSerLysCysAspVala 934
3302 TCAGAGAGATCTCGAGTAGACAGATTCGCGACGCTGCTGTATGAG 3351
934 sPlySerValGlnProGluProPhePheHisLysValValHisSerGlu 950
3352 AGAAGTCAGTACAGCCAGAAATGATTTTCCAAAGAGTGTTTCATTC 3401

```



```

951 HisLeuAsnLeuValProGlnValGlnSerValGlnCysSerProGlnGlu 967
    |||||
3402 GACCTTGAACTTA.....GTCAGCAGCACTTCATTCCTCACCAAGAGA 3442
    |
967 userPheAlaPheArgSerHisSerHisLeuProPheLysAsnLysAsnL 984
    |
3443 ACCAATTCACCAATTCAGATCTCAGTCTGATTCACCAAAAACTAAGAGCA 3492
    |
984 ysaAsnSerLeuLeuIleGlyLeuSerThrGlyLeuPheAspAlaAsnAsn 1000
    |||||
3493 AGAATTCCTTACGATTCGACCTTCACCTGCTGTTGATGACCAACAT 3542
    |||||
1001 ProlyMetLeuArgThrCysSerLeuProAspLeuSerLysLeuPheAr 1017
    |||||
3543 CCAAAAGTCTGAGAGACCTGCTCAGTCCAGATCTTCCAAAGCTGTGAG 3592
    |||||
1017 GThrLeuMetAspValProThrValGlyAspValArgGlnAspAsnLeuG 1034
    |||||
3593 AACCTTATGGACCTTCCTCCACTGTGGGACCTTCATCAAGACAGCTCTTG 3642
    |||||
1034 LuIleAspGluIleLysAspGluAsnIleLysGlnGlyProSerAspSer 1050
    |||||
3643 AAATCGAGACCTGGAAGATGAACCAATTAAGAGAGGCTTCTGATTC 3692
    |||||
1051 GluAspIleValPheGlnGluThrAspThrAspLeuGlnGluLeuGlnAl 1067
    |||||
3693 GAAGACACTGTATTTGAAGAACTGACACAGATTTCAAGACCTTCAGGC 3742
    |||||
1067 aSerMetGlnGlnLeuLeuArgGlnGlnProGlyGlnGlySerGlnG 1084
    |||||
3743 CCAATGAGACGACCTGCTAGGGAGCAACAGGTGAGCAATGAGTGAAGG 3792
    |||||
1084 LuGlnGluSerValLeuLysAsnSerAspValGlnProThrAlaAsnGly 1100
    |||||
3793 AGCAAGACTGTCTTTAAAGCAAGCATGTGAGCAGACAGACAGAGAGG 3842
    |||||
1101 ThrAspValAlaAspGluAspAspAsnProSerSerGluSerAlaLeuAs 1117
    |||||
3843 ACAGATGGCCCGACGAGAGAGCAACCCACGACGCAAGAGAGAGAGG 3891
    |||||
1117 nGlnGluThrPheHisSerAspAsnSerAspGlyGluIleAlaSerGluCysG 1134
    |||||
3892 CGAGGATGGACCTCAGATATATAGCGCTGAGACCACTAGATGATGTG 3941
    |||||
1134 LuCysAspSerValPheAsnHisLeuGlnGluLeuArgLeuHisLeuGln 1150
    |||||
3942 AATATGACAGTGTCTTAAACATTTAGAGAACTAAAGACTTCACTGGAG 3991
    |||||
1151 GlnGluMetGlyPheGlnLysPhePheGlnValArgGlnLysIleLysAl 1167
    |||||
3992 CAAGAAATGGCTTTGAAAAGTCTTGAAGTATATGAGAAAGTAAAGGC 4041
    |||||
1167 aIleHisGluAspGluAspGluAsnIleGluIleCysSerLysIleValG 1184
    |||||
4042 TATTCATGAGATGAGATGAAGATTAATTAATTTGTTCAACATTAAGTTG 4091
    |||||
1184 InAsnIleLeuGlnLysGlnLysGlnHisLeuThrAlaLysIleLeuHis 1200
    |||||
4092 AGAATATTTTGGCAATGAGCACGACATCTCTATGCCCAAAATCTGCAT 4141
    |||||
1201 LeuValMetAlaAspGlyAlaLysArgGlnGlnLysAspAsnGlu 1214
    |||||
4142 TTAGTCATGCGCAGATGAGACCTATCAGAGATATATGATGAA 4183
    |||||
seq_name: /net/abs06/SIDS1/gcgdata/hold-sequences/geneseq-emb1/NA2001A.DAT.AAH17731
seq_documentation_block:
ID AAH17731 standard; cDNA: 2975 BP.
AC AAH17731:
XX
XX 26-JUN-2001 (first entry)

```

```

XX Human cDNA sequence SEQ ID NO:17341.
DE
XX Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.
KW
XX Homo sapiens.
XX BP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000: 2000EP-0116126.
XX
XX 29-JUL-1999: 99JP-0248036.
XX
XX 27-AUG-1999: 99JP-0300252.
XX
XX 11-JAN-2000: 2000JP-0118776.
XX
XX 02-MAY-2000: 2000JP-0183767.
XX
XX 09-JUN-2000: 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI: 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX
XX Claim 8; SEQ ID 17341; 2537PP + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises at least 15 nucleotides and the combination
XX of the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and
XX in gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH13632
XX represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
XX
XX Sequence 2975 BP; 894 A; 556 C; 562 G; 963 T; 0 other;
XX
XX alignment_scores:
XX Quality: 2012.00 Length: 387
XX Ratio: 5.19 Gaps: 0
XX Percent Similarity: 100.000 Percent Identity: 100.000
XX
XX alignment_block:
XX US-09-783-320-4 x AAH17731
XX
XX Align seg 1/1 to: AAH17731 from: 1 to: 2975
XX
XX 91 ArgIleAsnAlaGlnLysGlyValLeuPheGlnGluAspGlnIleLeuAs 107
XX |||||||
XX 3 CGAATTAATGCTCGAAGAGCGTTTGTTCAGAGAGATCAGATTTTGA 52

```

[illegible]

seq_name: /net/abs06/SIDS19/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AA286797

```
seq_documentation_block:
ID   AA286797 standard; DNA; 1846 BP.
XX
XX   AA286797:
XX
XX   17-APR-2000 (first entry)
XX
XX   Human protein kinase homologue coding sequence, PKH-6.
XX
XX   Protein kinase homologue; human; PKH; diagnosis; therapy; cancer; AIDS;
XX   autoimmune disorder; inflammatory disorder; reproductive defect; asthma;
XX   diabetes mellitus; infertility; ovulatory defect; endometriosis;
XX   polycystic ovary syndrome, ss.
XX
XX   Homo sapiens.
XX
XX   US6013455-A.
XX
XX   11-JAN-2000.
XX
XX   15-OCT-1998; 98US-0173581.
XX
XX   15-OCT-1998; 98US-0173581.
XX
XX   (INCY-) INCYTE PHARM INC.
XX
XX   Hillman JL, Yue H, Yang YT, Corley NC, Gorgone GA, Azimzai Y;
XX   Lu DM, Bandman O, Guejler KJ;
XX   WPI: 2000-136321/12.
XX
XX   P-PSDB: AAY6753.
XX
XX   Nucleic acids encoding a human protein kinase homolog useful for
XX   preventing, diagnosing and treating cancer, autoimmune/inflammatory
XX   disorders and reproductive defects -
XX
XX   Claim 6; Column 65-68; 38pp; English.
XX
XX   This sequence encodes a human protein kinase homolog (PKH) of the
XX   invention. The PKH sequences may be used in the prevention, treatment and
XX   diagnosis of the PKH sequences associated with inappropriate PKH expression such
XX   as cancers, autoimmune/inflammatory disorders and reproductive defects.
XX   They may be used to treat disorders associated with decreased PKH
XX   expression such as cancers (e.g. lymphoma, melanoma and cancers of the
XX   breast, lung and prostate), autoimmune/inflammatory disorders
XX   (e.g. AIDS, asthma and diabetes mellitus) and reproductive
XX   defects (e.g. infertility, ovulatory defects, endometriosis and
XX   polycystic ovary syndrome). The DNA may be administered to treat diseases
XX   by rectifying mutations or deletions in a patient's genome that affect
XX   the activity of PKH by expressing inactive proteins or to supplement the
XX   patient's own production of PKH polypeptides. Additionally, the DNA may be
XX   used to produce PKH, according to standard recombinant DNA methodology.
```

CC by inserting the nucleic acids into a host cell and culturing the cell to
 CC express the protein. Conversely, antisense nucleic acid molecules may be
 CC administered to down regulate PKH expression by binding with the cells
 CC own PKH genes and preventing their expression. The DNA, and antisense
 CC sequences may also be used as DNA probes in diagnostic assays to detect
 CC and quantitate the presence of similar nucleic acid sequences in samples,
 CC and hence which patients may be in need of restorative therapy. They may
 CC also be used to study the expression and function of PKH polypeptides and
 CC their role in metabolism. The PKH polypeptides may be used as antigens in
 CC the production of antibodies against PKH and in assays to identify
 CC modulators (agonists and antagonists) of PKH expression and activity. The
 CC anti-PKH antibodies and PKH antagonists may also be used to down regulate
 CC PKH expression and activity. The anti-PKH antibodies may also be used as
 CC diagnostic agents for detecting the presence of PKH polypeptides in
 CC samples.

XX Sequence 1846 BP; 611 A; 321 C; 369 G; 545 T; 0 other;

Alignment scores: Length: 438

Quality: 1959.00 Gaps: 2

Percent Similarity: 90.868 Percent Identity: 89.269

alignment_block:
 US-09-783-320-4 x AA286797 ..

Align seg 1/1 to: AA286797 from: 1 to: 1846

```

1 MetGluTyrValArgLeuGlnLysIleGlyGluGlySerPheGly 17
154 ATGGAGAAATGATGTTAGACTACAGAAAGATTGGAGAGGTTGATTTGGAAA 203
17 sAlaIleuValLysSerThrGluAspGlyArgGlnTyrValIleLysG 34
204 AGCATTCTTGTAAATCTACAGAAAGATGGCAGACATGATATATATAGG 253
34 IuIeSnlIeSerArgMetSerSerLysGluArgGluGluSerArgArg 50
254 AATTAACATCTCCAGAAATGTCAGTAAAGAAAGAAAGAAATCAAGAGG 303
51 GluValAlaValIleuValAsnMetLysHisProsnIleValGlnTyrAr 67
304 GAAGTTCAGATTTGGCAAAACATGAGACATCCAAATATGTCAGATAG 353
67 gGluSerPheGluGluAsnGlySerLeuTyrIleValMetAspTyrCysG 84
354 AGAATCATTTTGA..... 366
84 IuGlyGlyAspLeuPheLysArgIleAsnAlaGlnLysGlyValLeuPhe 100
366 ..... 366
101 GlnGluAspGlnIleLeuAspTrpPheValGlnIleCysLeuAlaLeuLys 117
367 .....GGAATTTTGGACTGGTGTGTACAGATATGTTGGCCCTGAA 407
117 sHisValHisAspArgLysIleLeuHisArgAspIleLysSerGlnAsnI 134
408 ACATGTCATGATAGAAAATTCCTCATCGAGACATTTAAATCTCAGAAC 457
134 IePheLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAla 150
458 TATTTTAACTAAAGATGAGACAGTACACCTTGAGATTTTGAATGCT 507
151 ArgValLeuAsnSerThrValGlnLeuAlaArgThrCysIleGlyThrPr 167
508 AGATTTCTTAATAGTACTGTAGAGCGGCTCGAACTTCATVAGGAGCC 557
167 cTyrTyrLeuSerProGluIleCysGluAsnLysProTyrIrsAsnLys 184
558 ATACTACTGTGCCTGAATCTGTGAAAACAAACACTTACATATATAAAA 607

```

```

184 eTasPILeTPAlaLeuGlyCysValLeuTyrGluLeuCysThrLeuLys 200
608 GTGACATTTGGCGCTGGGGGTGGTCTTTATGACACTGTACACTAAA 657
201 HisAlaPheGluAlaGlySerMetLysAsnLeuValIleLysIleLe 217
658 CATGCTTTGAGCTGACATATGAAAACCTGTAGCAGAAATATATTC 707
217 rGlySerPheProValSerLeuHisTyrSerTyrIrsAsnLys 234
708 TGGATCTTTCCACTGTGTCTTGTGCAATTTTCCTATGATCCCGCA 757
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
758 TCGTGTCTCAGCTTATTTAAAGAAATCTAGGATAGACATCAGTCA 807
251 SerIleLeuGluLysGlyPheIleAlaLysArgIleGlyLysPheLeu 267
808 TCCATATTTGGAGAAAGTTTATAGCCAAACGCATTTGAAAGTTTCT 857
267 rProGlnLeuIleAlaGluGluPheCysLeuLysThrPheSerLysPhe 284
858 TCCTCAGCTTATTCAGAGAAATTTGTCTTAAACATTTTGGAGTTTG 907
284 IYSerGlnProIleProAlaLysArgProAlaSerGlyGlnAsnSerIle 300
908 GATCACACGCTATACCACTAAAGACCACTTCAGAGACAAACCTGAT 957
301 SerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyrGly 317
958 TCTGTTATGCCGTCCACAAATTTACAAAGCTGCCGCTAAATATGAT 1007
317 eProLeuAlaTyrLysLysTyrGlyAspLysLysLeuHisGlyLysP 334
1008 ACCTTTAGCATTAAGAAATATGAGATTAATAATTCAGAAAGAAAC 1057
334 rOleuGlnLysHisLysGlnAlaHisGlnThrProGluLysArgValAsn 350
1058 CACTGCAAAACATTAACAGGCCCATCAAACTCCAGAGAAAGATGAT 1107
351 ThrGlyGluGluArgArgLysIleSerGluGluAlaAlaArgLysArg 367
1108 ACTGGAGAGAAAGAGAGAAATATCTGAGAGACAGCAAGAAAGAG 1157
367 gLeuGluPheIleGlyLysGlyLysGlnLysAspGlnIleIleSerL 384
1158 GCTGGAATTTATTTGAAAAGATTAAGAACGGTAGGATCAGATTAT 1207
384 eumetLysAlaGluGlnMetLysArgGlnGluLysGluArgLeuGluArg 400
1208 TAATGAAGCTGAACAAATGAAGCAAGCAAGCAAGCAAGCTTGCAAG 1257
401 IleAsnArgAlaArgGluGlnGlyTyrPargAsnValIleuSerAlaGly 417
1258 ATTAATATGGGCCAGCGAACACAGATGAGAAATGCTAAAGTCTG 1307
417 YSerGlyGluValLysAlaProPheLeuGlySerGlyGly.....ThrI 432
1308 AAGCGGGAAGTAAAGTAGGCAATTTTATACCAATATGTATACGTACA 1357
432 IeAlaProSerSer 436
1358 TTTTCCCTCCAGT 1371

```

seq_name: /net/abss06/SIDSL/9cgcdata/hold_geneseg/geneseqn-emb1/NM2001A.DAT:AA011848

seq_documentation_block:
 ID AAD11848 standard; CDNA: 1846 BP.
 XX AAD11848;
 AC
 XX
 DT 25-sep-2001 (first entry)
 XX

DE Human protein kinase homolog-6 (PKH-6) cDNA.

XX Human; protein kinase homolog-6; PKH-6; cytostatic; protein therapy;
 KW vaccine; immunosuppressive; antisclerotic; antiabortive; adenocarcinoma;
 KW Acquired immune deficiency syndrome; AIDS; melanoma; cancer; bone; liver;
 KW breast; autoimmune disorder; multiple sclerosis; drug screening; anaemia;
 KW Crohn's disease; ectopic pregnancy; tubal disease; inflammatory disorder;
 KW reproductive disorder; polycystic ovary syndrome; asthma; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 154..1191
 FT /tag= a
 FT /product= "Human protein kinase homolog-6 (PKH-6)"

XX US6264947-B1.

XX 24-JUL-2001.

XX 20-OCT-1999; 9905-0420915.

XX 15-OCT-1998; 9805-0173581.

XX (INCYTE GENOMICS INC.

XX Bandman O, Tang YF, Hillman JL, Yue H, Guegler KJ, Corley NC;
 PI Gorgone GA, Azimzai Y, Lu DAM;
 DR WPI: 2001-450728/48.
 DR P-PSDB: AAE06211.

PT Human protein kinase proteins and homologs, useful for preventing,
 PT diagnosing and treating cancers, autoimmune/inflammatory disorders and
 PT reproductive disorders -

XX Example; Column 67-68; 38pp; English.

XX The present cDNA sequence encodes human protein kinase homolog-6 (PKH-6).
 CC Human protein kinase homologs (PKH) and their cDNA molecules are used in
 CC the prevention, diagnosis and treatment of diseases associated with
 CC increased or decreased expression of PKH. Examples of such disorders
 CC include, cancer (e.g. adenocarcinoma, melanoma and bone, breast and
 CC liver cancer), autoimmune/inflammatory disorders (e.g. Acquired Immune
 CC deficiency Syndrome (AIDS), anaemia, asthma, Crohn's disease and
 CC multiple sclerosis) and reproductive disorders (e.g. tubal disease,
 CC ectopic pregnancy and polycystic ovary syndrome). PKH, its catalytic or
 CC immunogenic fragment are used for screening libraries of compounds in any
 CC of the drug screening techniques. PKH nucleic acids are used to generate
 CC hybridisation probes useful in mapping the naturally occurring genomic
 CC sequences. PKH are also used as antigens in the production of antibodies
 CC against protein kinases (PK) and in assays to identify modulators of PK
 CC expression and activity. PKH is also used in protein therapy.

XX Sequence 1846 BP; 611 A; 321 C; 369 G; 545 T; 0 other:

XX alignment_scores:

XX Quality: 1959.00 Length: 438
 XX Ratio: 4.922 Gaps: 2
 XX Percent Similarity: 90.868 Percent Identity: 89.269

XX alignment_block:

XX US-09-783-320-4 x AAD11848

XX Align seg 1/1 to: AAD11848 from: 1 to: 1846

1 MetGluTYrValArgLeuGlnIleGlyGluGlySerPheGlyIle 17
 154 ATGGAGAGAGTGTAGACTACAGAGATTGACAGACGTTCAATTGGAAA 203
 17 salalleuvallyserthrGluaspGlyArgGlnTyrValIleGly 34

204 AGCCATTCTGTTAAATCTACAGAGATGGCAGACAGTATGTTATCAAG 253
 34 IuileasnlIeserArgMetSerSerLySGluArgGluGluSerArg 50
 254 AAATTAACATCTCAAGAAATGTCCAGTAAAGAAAGACAAATCAAGAGA 303
 51 GluValAlaValLeuAlaAsnMetLySHisProAsnIleValGlnTyr 67
 304 GAAGTTCAGATTTGGCAAAACATGAAGCATCCAAATATTTGCCAGTATG 353
 67 gGluSerPheGluGluAsnGlySerLeuTyrIleValMetAspTyrCys 84
 354 AGAATCATTTGAA..... 366
 84 IuGlyGlyAspLeuPheLyArgIleAsnAlaGlnLySGlyValLeuPhe 100
 366 366
 101 GlnGluAspGlnIleLeuAspTrpPheValGlnIleCysLeuAlaLeu 117
 367GCAATTTTGACCTGTTGTACACATATGTTTGCCCTGAA 407
 117 SHISValHisAspArgLySHisLeuHisArgAspIleLySHisSerGlnAsn 134
 408 ACATGTACATCATAGAAAATTTCTCATCGACATTAATCATCAGAACAA 457
 134 IepheLeuThrLyAspGlyThrValGlnLeuGlySHisPheGlyIleAla 150
 458 TATTTTAACATAAAGATGAGAACAGTACAACCTTGACATTTTGAAATGCT 507
 151 ArgValLeuAsnSerThrValGluLeuAlaArgThrCysIleGlyThrPr 167
 508 AGAATTTCTAATAGTACTAGTAGAGCTGGCTCGAAGCTTGATAGGACCC 557
 167 cTyrTyrLeuSerProGluIleCysGluAsnLySHisProTyrAsnAsnLyS 184
 558 ATACTACTGTGACCTGAAATCTGTGAACAAACCTTCAATATATATAA 607
 184 eraSpIleTrpPalaLeuGlyCysValLeuTyrGluLeuCysThrLeuLyS 200
 608 GTGACATTTGGGCTCTGGGTGTGCTTTATCAGCTGTGTACCTTAA 657
 201 HisAlaPheGluAlaGlySerMetLySHisLeuValLeuIleIleSe 217
 658 CATGCTTTTGAAGCTGGCAGTATGAAAACCTGCTCAAGATTAATATC 707
 217 rGlySerPheProProValSerLeuHisTyrSerTyrAspLeuArgSerL 234
 708 TGGATCTTTTCCACCTGTGCTTGCATATATTCCTATGATTCGCCAGTT 757
 234 euValSerGlnLeuPheLyArgAsnProArgAspArgProSerValAsn 250
 758 TGGTGCTCAGATTATTAAGAAATCTTAGGGATGACCATCAGTCAAC 807
 251 SerIleLeuGlnLySGlyPheIleAlaLyArgIleGluLyPheLeuSe 267
 808 TCCATATTGGAGAAAGTTTATATAGCCAAACGATTTGAAAATTTCTCTC 857
 267 rProGlnLeuIleAlaGluGluPheCysLeuLySHisPheSerLyPheG 284
 858 TCTCTACCTTATTCAGAAAGAAATTTGTCTAAAAACATTTTGAAGTTG 907
 284 LySerGlnProIleProAlaLyArgProAlaSerGlyGlnAsnSerIle 300
 908 GATCAGACCTATACAGCTAAAGAACAGCTTCAGAGCAAAACTCGATT 957
 301 SerValMetProAlaGlnLySHisIleThrLySHisProAlaAlaLySHis 317
 958 TCTGTATGCTGCTGCGAAATTTACAAAGCCGCGCGCTAATATATGAAAT 1007
 317 eProLeuAlaTyrLySHisTyrGlyLySHisPlySHisLeuHisGlyLySHis 334
 1008 ACCTTTAGCATATTAAGAAATATGAGATATAAAATATTACACGAAAGAAAC 1057

```
334 10LEuGlnLysHisLysGlnAlaHisGlnThrProGluLysArgValAsn 350
|||||
1058 CACGCGCAAAAACATTAACAGCCCAATCAACATCCAGAGAGAGAGTGTAT 1107
|||||
351 ThrGluGluGluArgArgLysLysSerGluGluAlaAlaArgLysArgArg 367
|||||
1108 ACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1157
|||||
367 GLeuGluPheIleGluLysGluLysLysSerGlnLysAspGlnIleIleSerL 384
|||||
1158 GCTGGAAATTTATTCAGAAAGATTAAGAGAGCGTGTGATCAGATTATAGCT 1207
|||||
384 eumetLysAlaGluGlnMetLysArgGlnGluLysGluLysArgLysArg 400
|||||
1208 TATGTAAAGCTGAAACAAATGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1257
|||||
401 ILeuAsnArgAlaArgGluGlnGluLysLysArgAsnValLeuSerAlaGly 417
|||||
1258 ATTAATAGAGCGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1307
|||||
417 ySerGluGluValLysAlaProPheLeuGlySerGlyGly.....ThrI 432
|||||
1308 AAGTGTGAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1357
|||||
432 LeuAlaProSerSer 436
|||||
1358 TTTTCCCTCCAGT 1371

seq_name: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAH06097

seq_documentation_block:
ID AAH06097 standard; cDNA; 876 BP.
AC AAH06097;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:2932.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EPI074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
XX
PR 27-AUG-1999; 99JP-0300253.
XX
PR 11-JAN-2000; 2000JP-0118776.
XX
PR 02-MAY-2000; 2000JP-0183767.
XX
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI: 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 1: SEQ ID 2932; 2537bp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
```

```
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences. AAH92446 to
CC AAH95893 represent human amino acid sequences, and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ
Sequence 876 BP: 307 A; 160 C; 159 G; 239 T; 11 other:

alignment_scores:
Quality: 1218.50 Length: 282
Ratio: 4.598 Gaps: 5
Percent Similarity: 93.972 Percent Identity: 89.716

alignment_block:
US-09-783-320-4 x AAH06097
Align seg 1/1 to: AAH06097 from: 1 to: 876

91 ArgIleAsnAlaGlnLysGlyValLeuPheGlnLysAspGlnIleLeuAs 107
|||||
3 CGAATAAATGCTCAGAAAGCGCTTTGTTTCAAGAGATCAGATTGTGGA 52
|||||
107 pTrpPheValGlnIleCysLeuAlaLeuLysHisValHisAspArgLysI 124
|||||
53 CTGCTTTGACAGATATGTTTGCCCTGAAACATGATACATGATGAAAAA 102
|||||
124 LeuLysHisArgAspIleLysSerGlnAsnIlePheLeuThrLysAspGly 140
|||||
103 TTCTTCATCGAGACATTAATCTCAGAAACATATTTTAACTAAAGATGGA 152
|||||
141 ThrValGlnLeuGlyAspPheGlyIleAlaArgValLeuAsnSerThrVa 157
|||||
153 ACAGTACAACTTGAGATTGGAATGCTAGAGTCTTAATTAAGTACTGT 202
|||||
157 IGluleuAlaArgThrCysIleGlyThrProTyrTyrLeuSerProGluI 174
|||||
203 AGAGCTGCTCGAACTTGATAGGAGCCCATACACTTGTCCACCTGAAA 252
|||||
174 IeCysGluAsnLysProTyrAsnAsnLysSerAspIleTrrAlaLeuGly 190
|||||
253 TCTGTGAAAAACCAACCTTCAATTAATTAATAAGACATTTGGGCTGTGGG 302
|||||
191 CysValLeuTyrGluLeuCysThrLeuLysHisAlaPheGluAlaGlySe 207
|||||
303 TGtGTCTTTATGAGCTGTATACACTTAACATGCTTTTGAAGCTGGCAG 352
|||||
207 TmetLysAsnLeuValLeuLysIleIleSerGlySerPheProProValS 224
|||||
353 TATGAAAAACCTGCTACTGAAATATATATCTGATCTTTTCCACCTGTCT 402
|||||
224 erLeuHisLysSerTyrAspLeuArgSerLeuValSerGlnLeuPheLys 240
|||||
403 CTTTGCATTTATCTCTATGATCTCCGAGTTGGTGTGAGTATTTTAAA 452
|||||
241 ArgAsnProArgAspArgProSerValAsnSerIleLeuGluLysGlyPh 257
|||||
453 AGAATATCCTAGGAGATAGACATCATGATCACTTCATATTTGAGAAAGTTT 502
|||||
```

```

257 eilealalysargllleglulyspheleuserprogluleulealaclug 274
|||||
503 tvtacccaacacgattgaaaggttctctccacacttaattgcacag 552
|||||
274 luphacysleulysrhpheuserlyspheglserglinproleprola 290
|||||
553 aattttgcttaaaaaacattttcgaaatttgatcacagccatpaccagct 602
|||||
291 lysargproalaeserglylnasserlleservalemetproalaglnly 307
|||||
603 aaaaaacgagcttcangacaaaacgcnatttctgtatgcctgctcaaa 652
|||||
653 aatttac.aagcctgccgcttaaatgatnctttagcattatagaaa 701
|||||
324 tyrglyasr.lyslyslau.hisglulysprouleuglnlyslsly 339
|||||
702 tatgganatttataaaatttaccnnaaaagAACACACCTGCCTAAACNTAAA 751
|||||
340 glualahsaglthrpro..glulysargvalasnthrlygluluar 355
|||||
752 ccngcgccctccacactcccaaaaaaaattgaaattctggcgaanaaaaaa 801
|||||
356 ..arglyslieserglulualalaarglysargarg 367
|||||
802 gggaggaataatttctgaagnaaaccccccgaanaaaaaa 840
|||||
seq_name: /net/abs06/SIDSI/gcgdata/hold1-geneseq/geneseg-emb1/NA2001A.DAT:AA506745
seq_documentation_block:
ID AA506745 standard; cDNA; 1947 BP.
XX
AC AA506745:
XX
DT 12-SEP-2001 (first entry)
XX
DE Polynucleotide sequence encoding human protein kinase #45.
XX
KW Human: protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200138503-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000MO-US32085.
XX
PR 24-NOV-1999; 99US-0167482.
XX
PA (SUGEN-) SUGEN INC.
XX
PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R.
PI Flanagan P, Clarry D;
XX
WP1: 2001-343950/36.
XX
DR P-PSDB; AAU03545.
XX
NU Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections -
XX
XX
XX Example 1: Figure 1: 433bp; English.
XX
XX AA506701-AA506757 encode for novel human protein kinases #1-57. The
XX novel protein kinases have been identified as members of the tyrosine
XX or serine/threonine kinase (PTK and STK) families. The polynucleotides
XX encoding protein kinases and the polypeptides may be used in the
XX prevention, diagnosis and treatment of diseases associated with

```

```

CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of hematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be
CC used for gene therapy and as DNA probes in diagnostic assays.
CC The protein kinase polypeptides may be used as antigens in the production
CC of antibodies against the protein kinases and in assays to identify
CC modulators of protein kinase expression and activity.
XX
SO Sequence 1947 BP; 638 A; 357 C; 455 G; 497 T; 0 other;

alignment_scores:
Quality: 1025.00 Length: 818
Ratio: 2.293 Gaps: 25
Percent Similarity: 54.645 Percent Identity: 32.763

alignment_block:
US-09-783-320-4 x AA506745 ..

Align seg 1/1 to: AA506745 from: 1 to: 1947
1 Metglulysrlyrvalargleuglnlylleglulglulysrhpheglyly 17
|||||
1 ATGCAATAGTACGATGATGATTAAAGCCATCGGCAAGTGCCTCGGGCA 50
|||||
17 salatleuvalylserrhrghlualspgllyargglnlyrvaliilelysg 34
|||||
51 AGCATACTTAGCTAAAGCAATACAGATAGCAAGACGTGTCATATAAAG 100
|||||
34 lulleasnliieserargmetserserlysglnargluluglserrargarg 50
|||||
101 AGATCAATTTTGAANAAGATGCCATACACAAANAAGACCTTCAANAANA 150
|||||
51 glualalavalaleualasmetlyshisproasnlilealglntyrar 67
|||||
151 GAGCTGATTTCTTGTGAAAGATGAANAACATCCACATTTGACCTTCT 200
|||||
67 gluserhpheglulualunglyserleutyrlilevalmetasprlyrcysg 84
|||||
201 CAATTCAATTCACAGAAATGCGAGCGCTTATTGTAATGAATATTGTG 250
|||||
84 luclyglyaspheupheleusergileasnlaaglnlysglyvalleuphe 100
|||||
251 ATGAGAGGGGATCTCATGAAAGATCAATAGACACAGCGGTGTGTATT 300
|||||
101 glngluaspgrlnlileuasprrprrprrprrprrprrprrprrprrprr 117
|||||
301 AGTGAGATGATGATCTCGGTGGTGTGTGTGTGTGTGTGTGTGTGTGT 350
|||||
117 shisvalhisasparglysileuuhisargasprrprrprrprrprrprr 134
|||||
351 ACATATTCATGACGAGAAATATTACACAGGACATATAAGCTCAGACACA 400
|||||
134 lepheleuthrlyaspelgylthrval...glnleuglysrprrprrprr 149
|||||
401 TTTTCTTAGCAAGACGAAATGCTGCAAGCTTGGGACCTTGGATATA 450
|||||
150 AlaargvalleuasnserrhrvalgluleualaarglthrCyslileglyth 166
|||||
451 GCAAGAGCTCTGAATTAATTCATGAACTTGCAGAACTTGAATGGAAAC 500
|||||
166 rprolytyrleuserproglulilecysguasnllysprrprrprrprrprr 183
|||||
501 ACCTTACACTGCTGCCACAGATCTGTGAGATTAACCTTCAACAAATA 550
|||||
183 ysserasprrprrprrprrprrprrprrprrprrprrprrprrprrprr 199
|||||
551 AAACGATATTGCTCTCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 600
|||||

```

200 LysHisAlaPheGluAlaGlySerMetLysAsnLeuValLeuIleIle 216
 601 AAACATCTCTTTGAGGGTAAACAACTTACACACCTGCTTCTGAAGATTGG 650
 216 eSerLysSerPheProValSerLeuHisTyrSerTyrAspLeuArgS 233
 651 TCAACACATTTTGGCCCCAATATCTCCGGGGTCTTCTCGAGACCTCATTT 700
 233 eLeuValSerGlnLeuPheLysArgAsnProArgAspArgProSerVal 249
 701 CCTTATATCTCAGCTCTTCAAGTATCTCTGAGACCGACCATCATATA 750
 250 AsnSerIleLeuGlnLysGlyPheIleAlaLysArgIleGlnLysPhe 266
 751 AATTCATTTTCAAAAGGCCCTTTTAGAGAACTTATCTCCAAATATATT 800
 266 uSerProGln.....LeuIleAlaGlnGluP 275
 801 GACTCTGAGGTAAAGTATTGAGCGTACTGTTGATTTTGGCAGAGATT 850
 275 heCysLeuLysThrPheSerLys.....PheGlySerGlnProIle 288
 851 TGGGTTGCGAGTCTTGACAGCTGTGTTGGTTTAGGTGTGAGCCACTG 900
 289 ...ProAla.....LysArgProAlaSerGlnAsnSe 299
 901 CACCCAGCCCTGATTATGTTTTTAAACATCCCTCTGTTTCTTCA 950
 299 rIleSerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyrG 316
 951 GATTAATAATGATAGAAGAACCCCAAAATGCTGCTGTGACATATATG 1000
 316 LylLeuProLeuAlaTyrLysLysTyrGlyAspLysLysLeuHisGlnLys 332
 1001 ATTATTATTATGCTCAACTTGATATGCTAGAGAGAGAGCCCA..... 1044
 333 LysProLeuGlnLysHisLysGlnAlaHisGlnThrProGlnLysArgVal 349
 1045 AAACCA.....AGTTATCACCCCTATCTCTCAAGAA..... 1074
 349 LAsnThrGlyGlnGluArgArgLysIleSerGlnGlnAlaAlaArgLysA 366
 1075 AATCTGGAGTGTGAG..... 1089
 366 rGArgLeuGlnPheIleGlnLysGlnLysLysGlnLysAspGlnIleIle 382
 1090GATTACGGCTCAGAAACAGGCAATGTCATCCCA..... 1125
 383 SerLeuMetLysAlaGlnLysMetLysArgGlnGlnLysGlnArgLeuG 399
 1126 AGTAAATGGCTGCTGAGTACCTTCAGAGAAAATTGGAAGCTCAACAATA 1175
 399 uArgIleAsnArgAlaArgGlnGlnLysTyrPArgAsnValLeuSerAlaG 416
 1176 TAACTTGAAGTGAAGAACAAATTTGGTCTCT..... 1209
 416 LysGlySerGlnLysValLysAlaProPheLeuGlySerGlyLysThrIle 432
 1209 1209
 433 AlaProSerSerPheSerSerArgGlyGlnTyrGlnHisTyrHisAlaI 449
 1210 ...CCATCTTCTGCCGAGCCAAT..... 1230
 449 ePheAspGlnMetGlnGlnGlnAlaGlnAlaAspAsnGlnAlaLysTrpL 466
 1231 TACAGACGAGACAGAGAGCTAAGAGTATGAGAGAGAGCTAGATTCC 1279
 466 ySarGlnIleTyrGlyArgGlyLeuProGlnArgGlnLysGlnGlnLeu 482
 1280 AGGAG.....CTGCCATTTTGAAGAAAAGCAATGAG 1311

483 AlaValGluArgAlaLysGlnValGlnGlnPheLeuGlnArgLysArgI 499
 1312 GAACAGGAATATTGGAACCACTTAGAGGAA..... 1341
 499 uAlaMetGlnAsnLysAlaArgAlaGlnGlnHisMetValTyrLeuAlaA 516
 1341 1341
 516 rGLeuArgGlnIleArgLeuGlnAsnPheAsnGlnArgGlnGlnIleLys 532
 1342ATAGCCCAACAGTACCAACATGACATGAAAGAAATTAAG 1380
 533 AlaLysLeuArgGlyGlnLysLysGlnAlaAsnHisSerGlnLysGlnI 549
 1381 AAGAAATGGGAGAGAAACAGAGGACTT..... 1410
 549 uGlySerGlnGlnAlaAspMetArgArgLysLysIleGlySerLeuLysA 566
 1411GAAAAAGACTTGAAACAAATGAGGCTTCAGAACACAAAG. 1449
 566 LAsnAlaAsnAlaArgAlaAlaValLeuLysGlnGlnLeuGlnArgLys 582
 1450GAAAGTAAA 1458
 583 ArgLysGlnAlaTyrGlnArgGlnLysLysValTrpGlnGlnHisLeuVal 599
 1459 AATCCAGAA.....CAGAAATTTAAGC 1481
 599 LAlaLysGlyValLysSerSerAspValSerProProLeuGlyGlnHisG 616
 1482 TAAAGAGGGGTAAATT..... 1500
 616 LuthrGlyGlySerProSerLysGlnGlnMetArgSerValIleSerVal 632
 1500 1500
 633 ThrSerAlaLeuLysGlnValGlyValAspSerSerLeuThrPArg 649
 1501GAAATATTTAGCAAAATGATTTCTGAT..... 1530
 649 gGluThrSerGlnLysMetGlnLysThrAsnAlaIleSerSerLysA 666
 1530 1530
 666 rGlnIleLeuArgArgLeuAsnGlnLysAsnLeuLysAlaGlnGlnAspG 682
 1531GAAACATCTCTCCAGAGAGAGAGCA 1557
 683 LysGlyMetGlnAsnLeuSerAspThrPheGln.....LeAs 695
 1558 ATGATATACCAAAATGAACCTTGACCTTGAAGATGCGCATGAATTTAA 1607
 695 nValHisGlnAspAlaLysGlnHis.....G 704
 1608 GGAATATGAATGCTAAGAGCATGAGATTTACAGACAAAGCATTTTG 1657
 704 LysLys.....GluLysSerValSerSer..... 711
 1658 AAAAATCTCACTGCCAGAGAGGAGGTTTCCAGCAGACTGTAGCTGCT 1707
 712AspArgLysLysTrpGlnAlaGlyLys...GlnLeuValIlePr 725
 1708 GTGGAAACAGAGAGAGAGCTGAGTGAAGAGAGAGAGAGAGAGAGAGAG 1757
 725 uLeuAspGlnLeuThrLeuAspThrSerPheSerThrThrGlnAlaGlnHisT 742
 1758 GATGATGGCAGCTGGCGAGCATCACTCTCACTGCCAGAGGGGCGCTGACA 1807
 742 hValGlyGlnValIleLysLeuGlyProAsnGlySerProArgArgAla 758
 1808 AT...GGCCAAAGTTATTTGATTAAGGATTCAGAGAAACAGAGAAACAG 1854
 759 TrpGlyLysSerProThrAspSerValLeuLysIleLeuGlyGlnAlaG 775

• Wed May 15 14:14:00 2002

649 rggju 650
|||||
848 GGGAA 852

us-09-783-320-4.rng

THIS PAGE BLANK (USPTO)


```

134 IepheLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAla 150
151 ArgValIleAsnSerThrValGlnLeuAlaArgThrCysIleGlyThrPr 167
451 AGAGTTCTTAATAGTACTGTAGACTGCTGCAGAACTTCATAGGAGCC 500
167 CTTTCTTCTTAAATAGTACTGTAGACTGCTGCAGAACTTCATAGGAGCC 500
501 ATAGACTTGTACCTGAAATCTGTGAAACCAACCTTACATTAATATTA 550
164 eAspIleThrAlaLeuGlyCysValLeuThrGlyLeuGlyThrLeuLys 200
551 GTGACATTTGGGCTCTGGGGTGTCTCTTATGAGCTGTGTACACTTAA 600
201 HisAlaPheGlnAlaGlySerMetLysAsnLeuValLeuLysIleIleSe 217
601 CATGCTTTTGAAGCTGGCAGATGAAAAACCTGTACTGAAGATTAATATC 650
217 rGlySerPheProValSerLeuHisTyrSerTyrAspLeuArgSerL 234
651 TGGATCTTTTCCACCTGTCTCTTTCATTTCTTATGATCTCCGACGTT 700
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
701 TGGTCTCTCAGTTATTTAAAGAAATCTAGGAGTACCATCAGTCACAC 750
251 SerIleLeuGlnLysGlyPheIleAlaLysArgIleGlnLysPheLeuSe 267
751 TCATATATGGAGAAAGGTTTATAGCCAAACGATTTGAAAAAGTTCTCTC 800
267 rProGlnLeuIleAlaGlnLysPheCysLeuLysThrPheSerLysPheG 284
801 TCCTAGCTTATTCAGAGAAATTTTCTCTTAAACATTTTCCAACTTGG 850
284 LysSerGlnProIleProAlaLysArgProAlaSerGlyGlnAsnSerIle 300
851 GATTCACAGCTTATACAGCTTAAAGAACCTTCAGAGCAAAACCTGCAAT 900
301 SerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyrGlyL 317
901 TCTGTATAGCCTGCTCAGAAAAATTAACAAGCTGCCCTTAATATGGAAT 950
317 eProLeuAlaTyrLysLysTyrGlyAspLysLysLeuHisGlnLysLysP 334
951 ACCTTTACCATATTAAGAAATGTGAGAGATTAATAAATTTACAGCAAAAGAAC 1000
334 rOleuGlnLysHisLysGlnAlaHisGlnThrProGlnLysArgValAsn 350
1001 CACTGCAAAACCATTAACAGAGCCCATCAACCTCCAGAGAAAGAGTGAAT 1050
351 ThrGlyGlnLysArgArgLysIleSerGlnGlnAlaAlaArgLysArgR 367
1051 ACTGAGAAAGAAAGAGAGAAATATCTGAGSAAACAGCAAGAAAGAGAG 1100
367 GlnLeuGlnPheIleGlnLysGlnLysLysGlnLysAspGlnIleIleSerL 384
1101 GCTGAGATTTTATGAAAAAGAAAGAAACAAAGAGATCAGATTAATTAAGTT 1150
384 euMetLysAlaGlnLysMetLysArgGlnLysGlnLysGlnLysGlnLys 400
1151 TAAATGAAGCTGAACCAATGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1200
401 IleAsnArgAlaArgGlnGlnLysTyrPArgAsnValIleuSerAlaGly 417
1201 ATAAATAGGCCAGGAGAAAGAGATGAGAAATGTCTAAGTCTGCTGCTG 1250
417 ySerGlyGlnValLysAlaProPheLeuGlnLysGlyGlyThrIleAlaP 434
1251 AAGTGTGAAGTAAAGGCTCTTTTCTGGGAGAGTGGAGGACATATAGCTC 1300
434 rOserSerPheSerSerArgGlyGlnLysTyrGlnHisTyrHisAlaIlePhe 450
1301 CATATCTTTTCTCTCCGAGGACAGTATGACATTAACATCCATTTT 1350
451 AspIleMetGlnGlnLysArgAlaGlnAsnGlnAlaLysTyrPArg 467
1351 GACCAATTCAGCAGCAACAAAGAGAGAGATTAATGAAGCTTAATGGAAAA 1400
467 gGlnIleTyrGlyArgGlyLeuProGlnLysGlnLysGlnLysAla 484
1401 AGAAATATATGCTCGAGCTCTTCCAAAGCAAAAGGAGGAGCTAGCTG 1450
484 alGlnArgAlaLysGlnValGlnLysPheLeuGlnArgLysArgGlnAla 500
1451 TGAAGAGCTTAACAGATGAGAGAGTCTCTGAGAGAAAGGAGAGCT 1500
501 MetGlnAsnLysAlaArgAlaGlnLysHisMetValTyrLeuAlaArgLe 517
1501 ATGAGATTAATTAAGCTCTGAGCCGAGAGACATTAATGTTTATCTGGCAAGCT 1550
517 uArgGlnIleArgLeuGlnAsnPheAsnGlnLysArgGlnLysAlaL 534
1551 GAGCAAAATTAAGCTACAGAAATTTCAATGAGCCCAACAGATTAAGCCA 1600
534 yLysLeuArgGlyGlnLysLysGlnAlaAsnHisSerGlnGlnLysGln 550
1601 AACTTCTGTGTAAGAAAGAAAGATTAATCATTTCTGAAGGACAAAGAGA 1650
551 SerGlnGlnAlaAspMetArgArgLysIleGlnSerLeuLysAlaHis 567
1651 AGTGAAGAGCTGACATGAGGGCGCAAAAAATCGAATCATCTGAAGGCCA 1700
567 sAlaAsnAlaArgAlaAlaValLeuLysGlnLysGlnLysArgLysArgL 584
1701 TGCATTAAGCAGCTGCTGCTGTACTTAAAGAACATAGACCAAGAGADA 1750
584 ySGlnAlaTyrGlnArgGlnLysLysValTyrGlnGlnHisLeuValAla 600
1751 AGGAGGCTTATGAGAGAGAAAAAGTGTGGAAAGCATTGTGGTGGCT 1800
601 LysGlnLysSerSerAspValSerProLeuGlnHisGlnLys 617
1801 AAAGGATTAAGAGTCTGATGTTTCTCCACCTTTGGGACACAGCAAGAAC 1850
617 rGlyGlySerProSerLysGlnLysMetArgSerValIleSerValIthr 634
1851 AGTGGCTCTCCATCAACAGCAAGATGAGATCTGTTATTTCTGTAACTT 1900
634 eAlaLeuLysGlnValGlyValAspSerSerLeuThrAspThrArgLys 650
1901 CAGCTTGAAGAAAGTTGGGGTGGACAGTAGTTTAACTGATACCCGGGAA 1950
651 ThrSerGlnLysMetGlnLysThrAsnAlaIleSerSerLysArgG 667
1951 ACTTCAGAAAGATGCAAAAAGCAACATGCTATTTCAAGTACGCGAGA 2000
667 uIleLeuArgArgLeuAsnGlnLysLeuLysAlaGlnLysAspGlnLysG 684
2001 AATTACTTCGAGATTAATGAATTAATCTTAAGCTCAAGAGAGTGAAGAAAG 2050
684 LysMetGlnAsnLeuSerAspThrPheGlnIleAsnValHisGlnLysAla 700
2051 GAATGAGCAATTCCTCTGATTAATCTTTGAGATTAATGTTTCAATGAATGCC 2100
701 LysGlnHisGlnLysGlnLysSerValIleSerSerAspArgLysTyrP 717
2101 AAAGAGCATGAAAAAGAAAAATCAGTTTCATCTGATCCGAGAAATGGAGA 2150
717 uAlaGlyGlnLeuValIleProLeuAspGlnLeuThrLeuAspThr 734
2151 GGCAGAGAGCTCACTTGTGATTTCTCTGATGAGTTAAACACTAGATTAAT 2200
734 ePheSerThrThrGlnArgHisThrValGlyGlnValIleLysLeuLys 750

```

2201 CCTCTCTACACTGAAGACATACAGTCGAGAAATTATTAACTAGCT 2250
 751 ProasnglyserProarararalarpilysSerProthraspSerVa 767
 2251 CCTAATGATCTCCAAAGAGAGCTGGGAGAAAGTCCGACAGATTCGT 2300
 767 lLeuylleleuglueglueglueglueglueglueglueglueglue 784
 2301 TCTAAGATCTTGAGAGAGCTGACACTGACAGAGATTCATTAAG 2350
 784 lueasrthrthrleasrsergluileserProgluglueglueglueglue 800
 2351 AAATACACTATTAGAGAGATTCCTCCGAGAGGAGAAAGATACAA 2400
 801 Proleuilethrleuglueglueglueglueglueglueglueglue 817
 2401 CCTTAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2450
 817 nProserAlallevalaspserProvalglueglueglueglueglue 834
 2451 CCCATCAGCTTATGTTGTTCTCTGTTGAGACAAAAGTCCGAGTTCA 2500
 834 ergluialeSerProglumetSerleuylleuglueglueglueglue 850
 2501 GTAGAGCATCTCCACAGATGTCATTAAGTGAAGAGAAATTTAGAG 2550
 851 Proasrleuylleuglueglueglueglueglueglueglueglue 867
 2551 CCTGATGATTGGAAACGAAATTCACAGAGCCAGAGTGAAGCAAA 2600
 867 sasrleuylleuglueglueglueglueglueglueglueglueglue 884
 2601 AAGTGAAGCTTGCCATGCTGATGATGATGATGATGATGATGATG 2650
 884 ysrglueglueglueglueglueglueglueglueglueglueglue 900
 2651 AAGAAACAAAGAGAACTCAGTCGAGATGATGATGATGATGATG 2700
 901 Gluevalserglueglueglueglueglueglueglueglueglue 917
 2701 GAAGTTCTGAGAGATGAGTCTCGAGTCTGAGTCTGAGTCTGAG 2750
 917 ehislleuglueglueglueglueglueglueglueglueglueglue 934
 2751 TCAATATAGAGCTCGAGAGCAATGATTCACAGCACTTAATGATG 2800
 934 splysrvalglueglueglueglueglueglueglueglueglue 950
 2801 ATAGTCTGTCACAGAGAGCAATGATTCATTAAGTCTGATTCGA 2850
 951 Hisleuasnleuvalprogluvalglueglueglueglueglueglue 967
 2851 CACTTGAAGCTTATGCTCCCTCAAGTCAATGATTCAGTCTGAC 2900
 967 uSerphealaphearSerHisSerHisSerHisSerHisSerHis 984
 2901 ATCTTTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 2950
 984 ysasrleuylleuglueglueglueglueglueglueglueglueglue 1000
 2951 AGAATTCCTGCTGATTCGATTCGATTCGATTCGATTCGATTCG 3000
 1001 Prolysmetleuylleuglueglueglueglueglueglueglueglue 1017
 3001 CCAAGATGTTAAGAGATTCCTCACTTCGATTCCTCAAGCTGTT 3050
 1017 gThrleuMetaspValProthrValglueglueglueglueglueglue 1034
 3051 AACCTTATGATGATTCCTCCAGCTGATGATGATGATGATGATG 3100
 1034 lueasrleuylleuglueglueglueglueglueglueglueglueglue 1050
 3101 AAATGATGAAATTTAAAGATGAGAAACATTAAGAGAGAGCTTC 3150

1051 Glueaspillevalpheuglueglueglueglueglueglueglueglue 1067
 3151 GAGAGCATTTGTTTGAAGAACTGACACAGATTTACAAAGAGCTCAG 3200
 1067 aserMetglueglueglueglueglueglueglueglueglueglueglue 1084
 3201 CTCGATGAGACAGTCTTGGAGACAACTGCTGACAGATTCATTAAG 3250
 1084 lueglueglueglueglueglueglueglueglueglueglueglueglue 1100
 3251 AAGAGAGTCACTGCTTGAAGACAGATGATGATGATGATGATGATG 3300
 1101 ThraspValalleaspglueglueglueglueglueglueglueglueglue 1117
 3301 ACAGATGTCGAGATGAGATGAGCAATCCAGATGAGAGAGAGAGAG 3350
 1117 nglueglueglueglueglueglueglueglueglueglueglueglue 1134
 3351 CGAGAGATGCGACATCAGATTAACAGTGTGCTGAATGCTAGTGA 3400
 1134 lueglueglueglueglueglueglueglueglueglueglueglueglue 1150
 3401 AATGCGATGATGCTTTTACCATTTAGAGAGAGAGAGAGAGAGAG 3450
 1151 Glueglueglueglueglueglueglueglueglueglueglueglueglue 1167
 3451 CAG 3500
 1167 alenHsglueglueglueglueglueglueglueglueglueglueglueglue 1184
 3501 TATTCATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATG 3550
 1184 lueasrleuylleuglueglueglueglueglueglueglueglueglueglueglue 1200
 3551 AAATATTTTGGAGAGATGAGATGAGATGAGATGAGATGAGATG 3600
 1201 LeuvalMetAlaspGlueglueglueglueglueglueglueglueglueglueglueglue 1214
 3601 TTAGTATGTCAGATGAGAGAGAGAGAGAGAGAGAGAGAGATG 3642

seq_name: gb_pat:AX224518

seq_documentation_block: 5426 bp DNA linear PAT 10-SEP-2001
 LOCUS AX224518
 DEFINITION Sequence 50 from Patent WO0161016.
 ACCESSION AX224518
 VERSION AX224518.1 GI:15554730

KEYWORDS
 SOURCE
 ORGANISM

human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
 1 (bases 1 to 5426)
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS
 Walke,D.W., Hu,X., Nepomichy,B., Turner,C.A. and Zambrowicz,B.

TITLE
 Human kinases and polynucleotides encoding the same

JOURNAL
 Patent: WO 0161016-A 50 23-AUG-2001;
 Lexicon Genetics Incorporated (US)

FEATURES
 source
 1..5426
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 1852 a 927 c 1163 g 1484 t
 ORIGIN

alignment_scores:

Quality: 6243.00 Length: 1214
 Ratio: 5.143 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-783-320-4 x AX224518

Align seg 1/1 to: AX224518 from: 1 to: 5426

```

1 MetGluysTYrValArgLeuGlnLysIleGlyGluGlySerPheGlyLys 17
117
534 ATGGAGAAAGTATGTTAGACTACAGAAAGATTGGAGAAAGTTTCATTGGAAA 563
17 sAlaIleuValLysSerThrGluAspGlyArgGlnTyrValIleLysG 34
584 AGCCATCTCTGTTAAATCTACAGAAAGTGGAGACATGATGATATCAGAG 633
34 LuIleAsnIleSerArgMetSerSerLysGluArgGluGluSerArgArg 50
634 AAATTAACATCTCAAGATGTCAGTAAAGAAAGAAAGAAATCAAGAGAGA 683
51 GluValAlaValIleuAlaAsnMetLysHisProAsnIleValGlnTyrAr 67
684 GAAGTTGCAGTATTCGCAAAACATGACATCCAAATTTGTCCAGTATAC 733
67 gGluSerPheGluGluAsnGlySerLeuTyrIleValMetAspTyrCysG 84
734 AGAATCATTTGAGAAAGTGGCTCTCTACATAGTATGATGATTAAGTCTG 783
84 LuGlyGlyAspLeuPheLysArgIleAsnAlaGlnLysGlyValIleuPhe 100
784 AGGAGGGGATCTGTTTAAAGCAATTAATGCTCGAAAGGCGTTTGTTT 833
101 GlnGluAspGlnIleLeuAspTyrPheValGlnIleCysLeuAlaLeuLys 117
834 CAAGAGGATCAGATTTTGTGACTGCTTGTACAGATATGTTGGCCCTGAA 883
117 sHisValHisAspArgLysIleLeuHisArgAspIleLysSerGlnAsnI 134
884 ACATCTACATGATGAGAAATTTCTTCATCGAGACATTAATCTCAGACACA 933
134 LePheLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAla 150
934 TATTTTAACTAAAGATGAGACAGACACTTGGAGATTTTGGAAATGCT 983
151 ArgValIleuAsnSerThrValGlnLeuAlaArgThrCysIleGlyThrPr 167
984 AGAGTTCTTAATAGTACTGAGCTGCTGCTGCAATCTGCATAGAGGACCC 1033
167 cTyTrpTyrLeuSerProGluIleCysGluAsnLysProTyrAsnAsnLys 184
1034 ATACTACTTGTCACCTGAAATCTGGAACAACAACTTACAAATATATATAA 1083
184 eAspIleTrpAlaLeuGlnCysValLeuTyrGluLeuCysThrLeuLys 200
1084 GTGACATTTGGCTCTGGGCTGTCTCTTATATGAGCTGTGTACACTTAAA 1133
201 HisAlaPheGluAlaGlySerMetLysAsnLeuValLeuLysIleIleSe 217
1134 CATGCTTTTGAAGCTGCGAGTATGAAAAACCTGTACTGAAAGATTAATC 1183
217 rGlySerPheProProValSerLeuHisTyrSerTyrAspLeuArgSerI 234
1184 TGGATCTTTCCACCTGCTGCTTGGCATATATCTATGATCTCGCAGCT 1233
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
1234 TGGTGTCTCAGTATTTAAAAAGAAATCTAGAGGATACACATCAGTAC 1283
251 SerIleLeuGluLysGlyPheIleAlaLysArgIleGluLysPheLeuSe 267
1284 TCCCATATTGAGAAAGTTTATATAGCCAAACGATTTGAAAAAGTTTCTCTC 1333
267 rProGlnLeuIleAlaGlnLysPheCysLeuLysThrPheSerLysPheG 284
1334 TCCCTAGCTTATTTGCAAGAAATTTGTCTAAAAACATTTTCCAGAGTTTG 1383
284 LysSerGlnProIleProAlaLysArgProAlaSerGlyGlnAsnSerIle 300
117
1384 GATCAGACCTATACAGCTAAAAAGCCAGCTTCAGGACAAACCTCGATT 1433
301 SerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyrGlyL 317
1434 TCTGTATATGCTCTCCAGAAATTTACAAAGCTCGCCGCTAAATTAATG 1483
317 eProLeuAlaTyrLysLysTyrGlyAspLysLysLeuHisGluLysLysP 334
1484 ACCTTTACATATATAGAAATATGAGATTAATAAAATTCACGAAAAAGAAC 1533
334 rGluGlnLysHisLysGlnAlaHisGlnThrProGluLysArgValAsn 350
1534 CACTGCAAAAACATTAACAGCCCATCAACCTCCAGAGAAAGAGAGTGAAT 1583
351 ThrGlyGluGluArgArgLysIleSerGluGluAlaAlaArgLysArgAr 367
1584 ACTGAGAAAGAAAGAGAAATATCTGAGAAAGCAGCAAGAAAGAAAGAG 1633
367 gLeuGluPheIleGluLysGluLysLysGlnLysAspGlnIleIleSerL 384
1634 GCTGGATTTATTTGAAAGAAAGAAACAAAGGATCAGATTAATTAAGTT 1683
384 euMetLysAlaGluGlnMetLysArgGlnGluLysGluArgLeuLysArg 400
1684 TAAATGAAGGCTGAACAAATGAAAGCAAGAAAGAAAGGTTGGAAAGA 1733
401 IleAsnArgAlaArgGluGlnGlyTrpArgAsnValLeuSerAlaGlyG 417
1734 ATAAATATGGGCCAGGCAAGAGATGAGAAATGTCTTAAGTGCAGTGG 1783
417 ySerGlyGluValLysAlaProPheLeuGlySerGlyGlyThrIleAlaP 434
1784 AAGTGGGGAAGTAAAGGCTCTTCTGGGAGTGAGGAGGACTATATAGCTC 1833
434 rSerSerPheSerSerArgGlyGlnTyrGlnHisTyrHisAlaIlePhe 450
1834 CATCATCTTTTCTTCGAGAGACATATGAACATTAACCATCCATTTT 1883
451 AspGlnMetGlnGlnGlnArgAlaGluAspAsnGluAlaLysTrpLysAr 467
1884 GACCAATTCGACGACAAAGACAGCAAGATTAATGAAGCTTAATGGAAGA 1933
467 gGluIleTyrGlyArgGlyLeuProGluArgGlnLysGlyGlnIleuAlaI 484
1934 AGAAATATATGTTGAGAGTCTTCAGAAAGGCAAAAGGCGACGCTAGCTG 1983
484 aGluArgAlaLysGlnValGluGluPheLeuGlnArgLysArgGluAla 500
1984 TAGAAAGAGCTAAACAAAGTGAAGAGTCTCTCAGCGAAACGGGAGACT 2033
501 MetGlnAsnLysAlaArgAlaGluGlnHisMetValTyrLeuAlaArgLe 517
2034 ATCGAAGATTAAGCTCGAGCCGAGAGACATATGTTTATCTGCGAAGCT 2083
517 uArgGlnIleArgLeuGlnAsnPheAsnGluArgGlnGlnIleLysAlaL 534
2084 GAGCAATTAAGACTACAGAAATTTCAATGAGCGCCAAAGATTAAGCCA 2133
534 yLeuArgGlyGlyLysLysGluAlaAsnHisSerGlnGluLysIleGlnG 550
2134 AACCTTCTGTGTAAGAAAGAAAGATATTCATTTCTGAAGACAAAGAGCA 2183
551 SerGluGluAlaAspMetArgArgLysLysIleGluSerLeuLysAlaHi 567
2184 AGTAAGAGGCTGACATGAGGCGCAAAATTCGAATTCATAAGGCCCA 2233
567 sAlaAsnAlaArgAlaAlaValLeuLysGluGlnLeuGluArgLysArgL 584
2234 TGCAAATGCAAGCTGCTGCTACTAAAGAACACATGACAGAACAAAGAGA 2283
584 ysgAlaIaTyrGluArgGluLysLysValTrpGluGlnHisLeuValAla 600
2284 AGGAGCTTATGAGAGAGAAAAAAGTGTGGAGAGACATTTGTGTGCT 2333

```

```

601 LysGlyValIysSerSerAspValSerProProlLeuGlnHisGluThr 617
|||||
2334 AAAGGAGTTAAAGATTCTGATGTTCTCCACTTTGGAGACGATGAAC 2333
|||||
617 rGlyGlySerProSerLysGlnMetArgSerValIleSerValThrS 634
|||||
2384 AGGAGGCTCTCCATCAACACACAGATGATCTGTATTTCTGTAAC 2433
|||||
634 eValLeuLysGluValIleValAspSerSerLeuThrAspThrArgGlu 650
|||||
2434 CAGCTTTGAAAGAGTTGGCGTGCAGATTTAACTGATACCGGGAA 2483
|||||
651 ThrSerGlnGlnMetGlnLysThrAsnAsnAlaIleSerSerLysArgG 667
|||||
2484 ACTTCAGAGAGATGCAAAAGACCAACATGCTATTTCAAGTAACG 2533
|||||
667 uIleLeuArgArgLeuAsnGlnLysAsnLysAlaGlnGluAspGlnLys 684
|||||
2534 AATCTTCGACAGTAAATGAAATCTTAAAGCTCAGAGATGAAAAAG 2583
|||||
684 LysMetGlnAsnLeuSerAspThrPheGlnIleAsnValHisGluAsp 700
|||||
2584 GAATGCAAAATCTCTGATACTTTTGACATTAATGTCATGAAATGCC 2633
|||||
701 LysGluHisGlyLysGlyLysSerValSerSerAspArgLysLysTrpG 717
|||||
2634 AAAGGAGCTGAAAAAGAAAAATCAGTTTCATCTGATCCAGAGAGTGG 2683
|||||
717 uAlaGlyGlyLysLeuValIleProlLeuAspGluLeuThrLeuAspThr 734
|||||
2684 GGCAGAGAGTCACCTTGATCTCTCGAGTGGATTACACATGATGATAC 2733
|||||
734 ePheSerThrThrGluArgHisThrValIleGlyValIleLysLeuGly 750
|||||
2734 CCTTCTTCACAACTGAAAGACATCACTGGGAGAAAGTTATTAATGAT 2783
|||||
751 ProAsnGlySerProArgArgAlaTrpGlyLysSerProThrAspSer 767
|||||
2784 CCTAATGATCTCCAGAGAGCCTGGGGAAAAAGTCCGACACATTTCT 2833
|||||
767 LysLysIleLeuGlyGlnIleGluLeuGlnLeuGlnLeuGlnLeuG 784
|||||
2834 TCTAAAGATCTTGGAGAGAGCTGAACATCAACTTCAGACGAACTAT 2883
|||||
784 LysAsnThrThrIleArgSerGluIleSerProGlnGlyLysLysLys 800
|||||
2884 AAAATACAACTATTGAGAGTGAATTTCTCCGGAAGGGAAAAAGTACA 2933
|||||
801 ProlLeuIleThrGlyLysLysValGlnCysIleSerHisGlnIleAs 817
|||||
2934 CCTTAAATTACGAGAAAAAAAGTACAAATGATTTTCACATGAAATTA 2983
|||||
817 nProSerAlaIleValAspSerProValGluThrLysSerProGlnPhe 834
|||||
2984 CCCATCAGCTATTGTTGATTCCTCGTTGAGACAAAAAGTCCCAATTC 3033
|||||
834 eGlnIleAsnSerProGlnMetSerLeuLysLysGlnGlnAsnLeuGln 850
|||||
3034 GTGAGGCATCTCCACAGATGTCATTGAAACGTGAAGAAATTAAGAAG 3083
|||||
851 ProAspAspLeuGlnThrGluIleLeuGlnLysProSerGlyThrAsn 867
|||||
3084 CCTGATGATTTGGAAAAAGAAATTTCTACAGAGCCAAAGTGAACAA 3133
|||||
867 sAspGlnSerLeuProCysThrIleThrAspValTrpIleSerGlnLys 884
|||||
3134 AGATGAGAGCTGCGCATGCATATTACTGATGTGGATTTAGTAGAGAA 3183
|||||
884 LysGluThrLysGluThrGlnSerAlaAspArgIleThrIleGlnGln 900
|||||
3184 AAGAACAAGAAAGAACTCAGTCGCGAGATAGATCACCATTCCAGAAAT 3233
|||||

901 GluValSerGlnAspGlyValSerSerThrValAspGlnLeuSerAsp 917
|||||
3234 GAAGTTCTGAAGATGAGATCTCGACTGAGCCAACTAGTACAT 3283
|||||
917 eHisIleGlnProGlyThrAsnAspSerGlnHisSerLysCysAspVal 934
|||||
3284 TCATRTAGAGCTGAGAACCAATGATTTCTGAGCAGCTCAATATGATGT 3333
|||||
934 sPylsSerValGlnProGluProPhePheHisLysValValHisSerGlu 950
|||||
3334 ATAATCTGTCACACCGAACCAATTTTCCATAGAGTGGTTCATTCTGA 3383
|||||
951 HisLeuAsnLeuValProGlnValGlnSerValGlnCysSerProGln 967
|||||
3384 CACTTGAACCTTAGCCCTCAGATTCATACAGTTCACTGATTCACGAGA 3433
|||||
967 uSerPheAlaPheArgSerHisSerHisLeuProProlLysAsnLysAs 984
|||||
3434 ATCCTTTGCATTTGATCTCAGTCTGCACTTACCAAAAAATTAATAACA 3483
|||||
984 LysAsnSerLeuLeuIleGlyLeuSerThrGlyLeuPheAspAlaAsn 1000
|||||
3484 AGAATTCCTGCTGATTTGGACTTTCACCTGCTGTTGATGCAACACAC 3533
|||||
1001 ProlLysMetLeuArgThrCysSerLeuProAspLeuSerLysLeuPhe 1017
|||||
3534 CCAAGATGTTAAGACATGTTCACTCCAGATCTCCAAAGCGTGTTCAG 3583
|||||
1017 gThrLeuMetAspValProThrValIleValAspValArgGlnAspAsn 1034
|||||
3584 AACCTTTAGATGTTCCACCGTAGAGATGTTGCTGACAGAACATCTTG 3633
|||||
1034 LysIleAspGluIleLysAspGlnAsnLysGlnGlyProSerAspSer 1050
|||||
3634 AATATGATGAATTTAAGATGAAAAATTAAGAGAGACCTCTGATTC 3683
|||||
1051 GluAspIleValPheGlnGlnThrAspThrAspLeuGlnIleLeuGln 1067
|||||
3684 GAAGACATTTGTTTGAGAAAGACAGACAGATTTTACAGAGCTGAGGC 3733
|||||
1067 aSerMetGlnGlnLeuLeuArgGlnGlnProGlnGlnGlySerGlnG 1084
|||||
3734 CTCGATGAGACAGTTACTTAGGAAACACCTGGTGAAGAAATACAGTGA 3783
|||||
1084 LysGlnIleSerValLeuLysAsnSerAspValGluProThrAlaAsnG 1100
|||||
3784 AAGAGAGTCAGCTTGAGAGACAGATGTTGAGACCAACTGCAAATGG 3833
|||||
1101 ThrAspValAlaAspGlnAspAsnProSerSerGlnSerAlaLeuAs 1117
|||||
3834 ACAGATGTGCGAGATGAAGATGACAACTCCAGTGTGAAAGTGGCTGA 3883
|||||
1117 nGlnGluTrpHisSerAspAsnSerAspGlyGlnIleAlaSerGluCys 1134
|||||
3884 CGAAGAAATGCACCTCAGATMACAGTGAATGGAATTTGCTAGTAATGT 3933
|||||
1134 LysCysAspSerValPheAsnHisLeuGlnGlnLeuArgLeuHisLeu 1150
|||||
3934 AATGCGATAGTGTCTTAAACCATTTAGAGAACTGAGACTTCATCTGAG 3983
|||||
1151 GlnGlnMetGlyPheGlnLysPhePheGlnValIleGlyLysIleLysAl 1167
|||||
3984 CAGGAATAGGCTTTGAAAAATTTCTTGAGGTTTATGAGAAATAAGGC 4033
|||||
1167 aIleHisGlnAspGlnAspGlnAsnIleGlnLysSerLysIleValG 1184
|||||
4034 TATTCATGAAGATGAAGATGAATATTTGAATTTGTTCCAAATAATAGT 4083
|||||
1184 LysAsnIleLeuGlnGlnHisGlnHisLeuThrValLysIleLeuHis 1200
|||||
4084 AAAATATTTTGGGAAGTGAACATCAGACATCTTTATGAGCCAAAGTTCT 4133
|||||
1201 LeuValMetAlaAspGlyAlaTrpGlnGlnLysAspAsnAspGlu 1214
|||||

```

|||||
4134 TTAGTCATGCGACGATGACCTACCAAGAGATTAATGATGAA 4175

seq_name: gb_pr:AB067488

LOCUS AB067488 5497 bp mRNA linear PRI 15-SEP-2001

DEFINITION Homo sapiens mRNA for KIAA1901 protein, partial cds.

AB067488

AB067488.1 GI:15620860

KEYWORDS Homo sapiens brain cDNA to mRNA, clone:h03635.

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE Nagase,T., Kikuno,R. and Ohara,O.

XXI. The complete sequences of 60 new cDNA clones from brain which

code for large proteins

JOURNAL DNA research : an international journal for rapid publication of

reports on genes and genomes. 8 (4), 179-187 (2001)

MDLINE 11572484

REFERENCE Ohara,O., Nagase,T. and Kikuno,R.

Submitted (27-JUL-2001) Osamu Ohara, Kazusa DNA Research Institute,

Department of Human Gene Research, 1532-3, Yana, Kisarazu, Chiba

292-0813, Japan (E-mail:cdnaif@kazusa.or.jp,

URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,

Fax:81-438-52-3914)

FEATURES

source

1..5497

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="h03635"

/issue="type-brain"

/note="vector:pBluescriptII SK plus"

472..4269

/gene="KIAA1901"

<472..4269

/note="Start codon is not identified."

/product="KIAA1901 protein"

/protein_id="BAB67794.1"

/gb_xref="GI:15620861"

/translation="RCLRRIMERYVRLQIGESFGKALVSTEDGRQYVIREINI

SHMSKERESEREYAVLANMKHPNIVYRESFENGSLTYVMYCGDLFRINMO

alignment_scores: Quality: 6201.00 Length: 1258

Ratio: 5.112 Gaps: 1

Percent Similarity: 96.423 Percent Identity: 96.343

alignment_block:

US-09-783-320-4 x AB067488

Align seg 1/1 to: AB067488 from: 1 to: 5497

1 MetGluLysTyrValArgLeuGlnLysIleGlyGluGlySerPheGly 17
|||||
493 ATGAGAGAGATGTGTACACAGAGATGGAGAGGTTGATGGAGAA 542
17 AlaIleLeuValLysSerThrGluAspGlyArgGlnTyrValIleLys 34
543 AGCATTCTGTTAAATCTACAGAGATGGACAGATGTTATCAAG 592
34 LuIleAsnIleSerArgMetSerSerLysGluArgGluLysSerArg 50
593 AAATTAACATCTCAGATGTCAGATGAAGAGAGAGAGAGAGAGAG 642
51 GluValAlaValIleuAlaAsnMetLysHisProAsnIleValGlnTyr 67
643 GAAATTCAGTATTCGCAACATGAGAGAGAGAGAGAGAGAGAGAG 692
67 GlnLysSerPheGluGluAsnGlySerLeuTyrIleValMetAspTyrCys 84
693 AGAATCATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 742
84 LucGlyAspLeuPheLysArgIleAsnAlaGlnLysGlyValLeuPhe 100
743 AGGAGAGAGATCTGTTTACCGAATATGCTCAGAGAGAGAGAGAG 792
101 GlnGluAspGlnIleLeuAspTrpPheValGlnIleCysLeuAlaLeu 117
793 CAAAGAGATCAGATTTGGACCTGTTTACAGATATATGTTGGCCCTGA 842
117 SHISValHisAspArgLysIleLeuHisArgAspIleLysSerGlnAsn 134
843 ACATGTCATCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 892
134 LePheLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAla 150
893 TATTTTAACATAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 942
151 ArgValLeuAsnSerThrValGlnLeuAlaArgThrCysIleGlyThrPr 167
943 AGACTTCTTAAATGATCTAGTAGAGCTGGCTCAGAACTGATAGGAGACC 992
167 CTTTTLysSerProGluIleCysGluAsnLysProGluAsnLys 184
993 ATCTACTCTGACCTGAATCTGTGAACCAACCTTCAATTAATAAA 1042
184 erAspIleTrpAlaLeuGlyCysValLeuTyrGluLeuCysThrLeuLys 200
1043 GTGACATTTGGCTCTGGGTGTCTCTTATGAGCTGTGACCTTAAA 1092
201 HisAlaPheGluAlaGlySerMetLysAsnLeuValLeuLysIleIleSe 217
1093 CATCTTTTAAAGCTGACATGAGAGAGAGAGAGAGAGAGAGAGAG 1142
217 GlnLysSerPheProValSerLeuHisTyrSerTyrAspLeuAlaGlySer 234
1143 TGGATCTTTTCCACCTGCTCTTTCATTAATTCATGATGATCCAGATT 1192
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
1193 TGGGTCTCAGTATTTTAAAGAGAAATCTAGGAGATGACATCATCATC 1242
251 SerIleLeuGlnLysGlyPheIleAlaLysArgIleGluLysPheLeuSe 267

BASE COUNT 1863 a 933 c 1183 g 1518 t

ORIGIN


```

1243 TCCATATTGGAGAAAGGTTTATAGCCAAACGATTTGAAAGTTCTCTC 1292
267 rPrGlnLeu1Lea1glnPheCysLeu1sThrPheSer1sPheG 284
1293 TCCCTACACTTATTCAGAGAAATTTTCTTAAACATTTTCGAAGTTG 1342
284 LysGlnPro1LePro1a1sArgPro1a1SerG1yGlnAsnSer1Le 300
1343 GATCAGACGCTATACCGTAAGACACGCTTCAGAGCAAAAACGCAATT 1392
301 SerValMetPro1a1gln1s1LeThr1sPro1a1a1a1s1TyG1y1 317
1393 TCTGTATGCTGTGCTGAGAAATTTACAAAGCCGCGCTAAATATGAAAT 1442
317 eProLeu1a1a1Ty1s1s1Ty1s1s1Ty1s1s1Ty1s1s1Ty1s1s 334
1443 ACCTTTACATATAGAAATATGAGATATAAAATTTACAGCAAAAAGAAC 1492
384 rOleuGln1s1s1s1s1s1s1s1s1s1s1s1s1s1s1s1s1s1s 350
1493 CACTGCAAAAACATTAACAGCCATCACTCCAGAGAAAGAGTGAAT 1542
351 ThrG1yGln1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 367
1543 ACTGAGAGAAAGAGGAAATATCTGAGAGAGCAGCAAGAAAGAGAG 1592
367 gLeuGlnPhe1LeG1u1sG1u1s1s1s1s1s1s1s1s1s1s1s1s 384
1593 GCTGGAATTTATTTGAAAAAGAAAAAGAAAGGATTCAGATTTATGATT 1642
384 eMet1s1s1s1s1s1s1s1s1s1s1s1s1s1s1s1s1s1s1s1s1 400
1643 TAAATGAAGCTGAACCAATGAAAGGCAAGAAAGGATTTGGAGAA 1692
401 IleAsnArg1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1 417
1693 ATAAATAGGCGCAGGAGAAAGATGAGAAATGCTGTAAGTCTGTG 1742
417 ySerG1yGln1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1 434
1743 AAGTGTGAAGTAAAGGCTCTTTCTGGGACGTGAGGAGCTATATGCT 1792
434 rOse1s1s1s1s1s1s1s1s1s1s1s1s1s1s1s1s1s1s1s1s1 450
1793 CATCATCTTTCTTCTGAGAGACATGACATTTACCATGCAATTTT 1842
451 AspGlnMetGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 467
1843 GACCAATATGACACAAAGAGAGAGATATGAGCTTAATGCAAAAG 1892
467 gGlu1Le1y1s1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1 476
1893 AGAAATATATGCTGAGGCTTCCAGAAAGAGAAATTCGCTGAGATT 1942
476 ..... 476
1943 GTCCAGAGATTTCCTTATGCGGCTGACAGTATCACATTTTCCGATCT 1992
476 ..... 476
1993 GATGATATTAGAAAACTTTGAAAGATTGAAGCGGTGCTAAACAAAG 2042
477 ..... 477
2043 CAAATGCAAAAGGCAAAAGGCGACCTAGCTGTAGAAAGAGCTTAAAC 2092
490 a1G1u1a1u1Phe1eug1n1a1rG1s1a1a1a1a1a1a1a1a1a1 506
2093 TAGAAGAGTTCTCGACGCAAAAGGGAAGCTATGCAAGATTAAGCTCG 2142
507 A1aG1uG1y1a1s1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1 523
|||||
2143 GCCAGAGACATATGCTTATCTGCAAGACTGAGCAAAATTAAGACTACA 2192
523 naSPheAsnGlnuArg1Gln1a1Le1ys1a1a1s1e1u1a1rG1yG1u1 540
2193 GAATTTCAATGAGCGCCACACAGATTAAGCCAAATCTGTGTGTAAGAA 2242
540 ySG1u1a1a1s1n1a1s1e1rG1u1G1nG1u1G1y1s1e1rG1u1a1a1 556
2243 AAGAAGCTTAATCATCTGAAGAGCAAGAAAGAGAGTCAAGAGCTACAT 2292
557 Arg1a1rG1y1s11eG1u1s1e1u1a1a1a1a1a1a1a1a1a1a1 573
2293 AGCGGCAAAAAATCGAATCACTGAAGCCCAATGCAAAATGACG1GCTGC 2342
573 aVal1Leu1sG1u1nG1eug1n1a1rG1s1a1a1a1a1a1a1a1a1a1 590
2343 TGTACTAAAGAACACACTAGAACGAAAGAGAGGCTTATGAGAGAG 2392
590 Lu1y1s1e1a1a1rPc1uG1uH1s1e1u1a1a1a1sG1y1a1a1s1e1 606
2393 AAAAAAAGTGTGGAGAGACATTTGGCTTAAAGAGTAAAGAGTTCT 2442
607 AspVal1SerPro1LeuG1yGln1a1sG1u1ThrG1yG1SerPro1Ser 623
2443 GATGTTTCTCCACCTTTGGGACAGCATGAAGAGCTGCTCCATCAAA 2492
623 sGlnGlnMet1a1rG1s1e1a11e1s1e1a1rP1s1e1a1a1e1u1sG1u1a1 640
2493 GCACACAGATGAGATCTGTTATTTCTGTAATCTCAGCTTTGAAAGAGT 2542
640 lYy1a1a1s1e1s1e1r1e1u1Thr1a1sP1h1a1rG1u1h1s1e1rG1u1 656
2543 GCGTGACAG1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 2592
657 Lys1ThrAsnAsn1a1a1e1s1e1s1e1s1a1a1e1u1a1a1a1a1a1 673
2593 AAGACCAACAAATGCTATTTCAAGTAGAGCGAAATACCTTGATGATTA 2642
673 nG1u1a1n1e1u1s1a1a1nG1u1a1sP1u1y1sG1y1e1tG1n1a1s1e1 690
2643 TGAATCTTAAAGCTCAAGAGAGATGAAGAAAGAGACAAATCTCTG 2692
690 sP1h1e1u1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1 706
2693 ATACTTTGAGATTAATATGTCATGAAAGTGCAGAAAGGCAATGAAG 2742
707 Lys1SerVal1Ser1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1 723
2743 AAATCAGTTTCATCTGATCGCAGAGAGGAGGAGAGAGCTCACTGT 2792
723 l1e1P1o1e1u1a1sP1u1e1u1Th1e1u1a1sP1h1e1u1a1a1a1a1 740
2793 GATTCCTGAGATGAGTTAAGCTAGATACATCCTTCTAACAACGTGAA 2842
740 rGh1s1Th1r1a1G1yG1u1a11e1y1s1e1u1G1yP1o1a1nG1y1s1e1 756
2843 GACATACAGTGGAGAGATTAATTAATTAAGTCTTAATGATCCAGAGA 2892
757 Arg1a1a1rPc1y1y1s1e1rP1h1a1sP1e1r1a1e1u1s11e1u1s1 773
2893 AGAGCCCTGGGGAGAGAGTCCAGAGATTTCTGTTAAAGATACCTGAGA 2942
773 u1a1a1u1eug1n1eug1n1Th1e1u1eug1n1a1n1Th1r11a1rG1s 790
2943 AGCTGAAC1TAACA1CTTCAACAGACATATTAAGAAATTAACACTTTGAA 2992
790 e1rG1u11e1s1e1rP1o1G1u1G1y1u1y1s1y1r1y1sP1o1e1u11e1h1 806
2993 GTGAGATTTCTCCGAAGGGGAGAAAGTCAAAACCTTATATTAACGGGAA 3042
807 Lys1ys1a1Gln1cys11e1s1e1h1sG1u11e1a1n1Pro1Ser1a1a1 823
3043 AAAAAAGTACAAATGATTTACATGATGAATTAACCCATCAGCTATTTGTA 3092

```

823 pSerProValIgluThrLysSerProGluIupheSerGluAlaSerProGlnM 840
 840 eSerLeuLysLeuGluLysLeuGluIupProAspPheLeuGluThr 856
 3093 TTCTCCTGTTGACAAAAGTCCCGAGTTCAGTGAAGCATCTCCACAGA 3142
 840 eSerLeuLysLeuGluLysLeuGluIupProAspPheLeuGluThr 856
 3143 TGTCTATGAACGTGAAGAAATTTAGACAGACCTGATGATTGGAAAC 3192
 857 GluIleuGlnIupProSerGluThrAsnLysAspGluSerLeuProcy 873
 3193 GAATTCACAGAGCCAGTGAACAAACAAAGATGAGCTTCCCATG 3242
 873 sThrIleThrAspValIupPheSerGluIupLysGluThrG 890
 3243 CACTTTCATGATGTGGATTAGTGAAGAAAGAAACAAAGGAACTC 3292
 890 lSerAlaAspArgIleThrIleGlnIuAsnGluValSerGluAspGly 906
 3293 AGTCGGCAGATAGATCACCATTCAGGAAATGAAGTTCTGAAGATGA 3342
 907 ValSerSerThrValAspGlnLeuSerAspIleHisIleGluProGly 923
 3343 GTCTGAGTACTGTGACCAACTAGTACATTCATATGAGCCTGGAAC 3392
 923 rAsnAspSerGlnHisSerLysCysAspValAspLysSerValGluPro 940
 3393 CAATGATTCACGACCTCTAATGTGATGATTAAGTCTGTCAACCG 3442
 940 lupProPhePheHisLysValValHisSerGlnHisLeuAsnLeuValPro 956
 3443 AACCATTTTCCATAGAGTGGTTCATTCGACACCTTGAACTTAACTCC 3492
 957 GluValGlnSerValGlnCysSerProGluIupSerPheAlaPheArgS 973
 3493 CAAGTTCATCATGTTCACTGTTCCACGAGAAATCCCTTCATTTGCATC 3542
 973 rHisSerHisLeuProPheLysAsnLysAsnLysSerLeuLeuIleG 990
 3543 TCACGCGATTTCACCAAAAATTAACAAAGAAATCTCTCTGATG 3592
 990 lLysSerThrGlyLeuPheAspAlaAsnAsnProLysMetLeuArgThr 1006
 3593 GACTTTCACAGTGTCTGTTGATGCAAAACCAACCAAGATTTAAGGACA 3642
 1007 CysSerLeuProAspLeuSerLysLeuPheArgThrLeuMetAspValPr 1023
 3643 TGTTCACCTCCAGATCTCTCAAGCTGTCAAGAACCTTATGATGTTCC 3692
 1023 oThrValGlyAspValArgGlnAspAsnLeuGluIleAspGluIleLysA 1040
 3693 CACCGTAGAGATGTGCTCAAGACATCTTGAATGATGAATGGAAG 3742
 1040 spGluAsnIleLysGluIupProSerAspSerGluIupLeuValPheGlu 1056
 3743 ATGAAACATTTAAAGAGACCTCTCATCTCAAGCATGTTGTTGA 3792
 1057 GluThrAspThrAspLeuGlnIupLeuGlnAlaSerMetGlnLeuLeu 1073
 3793 GAACTGACACGATTTACAGAGCTCAGACCTCGATGAGACAGATTA 3842
 1073 uArgGluIupProGluIupLysSerGluIupGluIupSerValLeu 1090
 3843 TAGGGAACACCTGTGAAGATATACATGAAGAAAGATCGTCTTGA 3892
 1090 yAsnSerAspValGluProThrAlaAsnGlyThrAspValAlaAspGlu 1106
 3893 AGAACAGATGTGAGCAACTGCAATGAGACAGATGTGCGACAGATGA 3942
 1107 AspAspAsnProSerSerGluSerAlaLeuAsnGluIupThrPheSerS 1123
 3943 GATGACATCCACAGACAGTGAAGTGCCTGACGAAGAAATGACATCAGA 3992

1123 pAsnSerAspGluIupLeuAlaSerGluCysGluCysAspSerValPhe 1140
 3993 TAACAGTATGAGTGAATTCAGTGAATGATGATGATGATGATGATGAT 4042
 1140 snHisLeuGluIupLeuArgLeuHisLeuGluIupGluIupMetGlyPheGlu 1156
 4043 ACCATTATGAGGAACAGTCACTTCATCTGACAGAGAAATGCGCTTGA 4092
 1157 LysPhePheGluValIupGluLysIleLysAlaIleHisGluAspGlu 1173
 4093 AAATTCCTTGAAGTTATGAGAAATTAAGCTATTCATGAGATGAGA 4142
 1173 pGluAsnIleGluIupLysSerLysIleValGlnAsnIleLeuGluAsnG 1190
 4143 TGAATAATTTGAATTTGTTCAAAATAGTCAAAATATTTTGGGAAATG 4192
 1190 lHisGlnHisLysGluValAlaLysIleLeuHisLeuValMetAlaAspGly 1206
 4193 AACATCAGCATCTTTATGCCAAGATTCATTAATGATGACAGATGA 4242
 1207 AlaTyrglnGluAspAsnAspGlu 1214
 4243 GCCTACCAAGAAATATGATGA 4266
 seq_name: gb_pr:HSMB00531
 seq_documentation_block:
 LOCUS HSMB00531 5511 bp mRNA linear PRI 18-FEB-2000
 DEFINITION Homo sapiens mRNA; cDNA DKFZp564L2416 (from clone DKFZp564L2416).
 ACCESSION AL050385
 VERSION AL050385.1 GI:4914588
 KEYWORDS
 SOURCE
 ORGANISM
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE
 1 (bases 1 to 5511)
 AUTHORS Koehler,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
 TITLE Direct SubMISSION
 JOURNAL Submitted (15-MAY-1999) MIPS, Am Kiofierspitz 18a, D-82152
 Martinsried, GERMANY
 COMMENT
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by BMFZ (Biomedical Research Center at the Charite,
 Berlin/Germany) within the cDNA sequencing consortium of the German
 Genome Project.
 This clone (DKFZp564L2416) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further
 information about the clone and the sequencing project is available
 at <http://www.mips.biochem.mpg.de/proj/cDNA/>.
 Location/Qualifiers
 1..5511
 source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKFZp564L2416"
 /clone_1bp="564 (synonym: hfb2). Vector pAMP1; host
 XL-2Blue; sites NotI + SalI"
 /dev_stage=" fetal"
 /issue_type=" brain"
 /map="4"
 join(534..2291,2293..4263)
 /gene="DKFZp564L2416"
 /note="Strong similarity to mouse protein kinase nek1
 The frame shift was determined manually"
 /product="hypothetical protein"
 join(534..2291,2293..4263)
 /gene="DKFZp564L2416"
 5476..5481
 polyA_signal
 polyA_site
 BASE COUNT 1883 a 943 c 1190 g 1495 t
 ORIGIN

alignment_scores: Length: 1243
 Quality: 6197.00 Gaps: 2
 Ratio: 5.109
 Percent Similarity: 97.586 Percent Identity: 97.586

alignment_block:
 US-09-783-320-4 x HSM800531 ..

Align seg 1/1 to: HSM800531 from: 1 to: 5511

```

1 MetGluuStyValArgLeuGlnLysIleGlyGluGlySerPheGly 17
534 ATGAGAGAGTATGTAGACTACAGAGATGGAGAGGTTCTTGGAAA 583
17 SAIAIAleuValLysSerThrGlnAspGlnArgGlnValIleLys 34
584 AGCCATTCTGTATTAATCTACAGAGATGCGACAGATGTTATCAAG 633
34 IuIleAsnIleSerArgMetSerSerLysGluArgGluGluSerArg 50
634 AAATTAACTCTCAAGATCTCCAGTAAAGAAAGCAAAATCAAGAGA 683
51 GluValAlaValLeuAlaAsnMetLysHisProAsnIleValGln 67
684 GAAAGTTCAGATTGGCAAAACATGAAACATCCAAATATTGTCAGTA 733
67 gGluSerPheGluGluAsnGlySerLeuTyrIleValMetAspTyr 84
734 AGAATCATTTGAGAGAAATGCTCTCTCTACATAGTAATGATTA 783
84 LuGlyAlaAspLeuPheLysArgIleAsnAlaGlnLysGlyValLeu 100
784 AGGGAGGGATCTGTTTAAGCAATTAATGCTCAGAAAGCGCTTTG 833
101 GlnGluAspGlnIleLeuAspTyrPheValGlnIleCysLeuAla 117
834 CAAGAGATCAAGATTTTGACTGCTGTGTACAGATAGTTGGCCCTGA 883
117 SHISValHisAspArgLysIleLeuHisArgAspIleLysSerGln 134
884 ACATGTACTGATGATAAAATTTCTTCATCGACATTAATCTCAGACA 933
134 IepheLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyIle 150
934 TATTTTAACTAAGATGAAAGATATTAATCTTGAGATTGGCAATTGC 983
151 ArgValLeuAsnSerThrValGlnLeuAlaArgThrCysIleGly 167
984 AGAGTCTTAATATGACTGTAGAGCTGCTGAACTTGCAAGGAGCC 1033
167 oTyrTyrLeuSerProGlnIleCysGluAsnLysProTyrAsnLys 184
1034 ATACTACTTGTCACCTGAATCTGGAACCAACCTTACATAATATAA 1083
184 eArgPheIleThrAlaLeuGlyCysValLeuTyrGlnLeuCysThr 200
1084 GTGACATTTGGGCTCGGGGTGTCTTATGAGCTGTGTACCTTAAA 1133
201 HisAlaPheGlnAlaGlySerMetLysAsnLeuValLeuLysIle 217
1134 CATGCTTTGAAGCTGCGAGTATGAAAAACCTGTAAGTAATAATC 1183
217 rGlySerPheProProValSerLeuHisTyrSerTyrAspLeuArg 234
1184 TGGATCTTTCCACCTGTGCTTTCATATATCTATGATCTCCGAGTT 1233
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerVal 250
1234 TGGGTCTCAAGTATTTAAAGAAATCCATGAGGATAGACATCACT 1283
251 SerLeuGlnLysGlyPheIleAlaLysArgIleGlyLysPheLeu 267

```

```

1284 TCCATATTGGAGAAAGTTTATAGCCAAACGATTTGAAAAATTCTCTC 1333
267 rProGlnLeuIleAlaGluGluPheCysLeuLysThrPheSerLysPhe 284
1334 TCTCTACCTATTATTCAGAGAAATTTGTCTAAAAACATTTCTGAA 1383
284 LysGlnProIleProAlaLysArgProAlaSerGlyGlnAsnSerIle 300
1384 GATCACAGCCTATACAGCTAAAGACAGCTTCAGAGCAAAACCTCAT 1433
301 SerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyrGly 317
1434 TCTGTATGCTGCTCGAATAATTAACAACTCCGCGCTTAATATNGAT 1483
317 eProLeuAlaTyrLysLysTyrGlyAspLysLysLysGlnLysLys 334
1484 ACCTTACATATATGAAATATGAGATTAATAATTAACGAAAGAAAC 1533
334 rOLeuGlnLysHisLysGlnAlaHisGlnThrProGluLysArgValAsn 350
1534 CACTGCAAAAACATTAACAGGCCCATCAACTCCAGAGAGAGAGTGA 1583
351 ThrGlyGluGluArgArgLysIleSerGluGluAlaAlaArgLysArg 367
1584 ACTGGAGAAAGAAAGAGAAATATCTGAGAGAGCAGAGAAAGAGAG 1633
367 gLeuGluPheIleGlyLysGlnLysLysGlnLysAspGlnIleIle 384
1634 GCTGGAATTTATGAAAAAGAAAGAAACAAAGAGATCAGATTATTA 1683
384 euMetLysAlaGluGluMetLysArgGlnGluLysGluArgLeuGlu 400
1684 TATTAAGAGCTGAAACAAATGAAAGGCAAGAAAGAAAGTTGCAAA 1733
401 IleAsnArgAlaArgGluGlnGlyThrArgAsnValLeuSerAlaGly 417
1734 ATTAATAGGCGCAGGAGAACAGATGGAATGTGTAATGCTCTGCTG 1783
417 ySerGlyGluValLysAlaProPheLeuGlySerGlyGlyThrIleAla 434
1784 AAGTGTGAAGTAAGAGCTCTTCTTGCGCAGTGGAGGACATATAGCT 1833
434 rSerSerPheSerSerArgLysGlnTyrGlnHisTyrHisAlaIlePhe 450
1834 CATCATCTTTCTCTCCAGAGACGTATGAACTATTCATGCCATTTT 1883
451 AspGlnMetGlnGlnArgAlaGluAspAsnGluAlaLysTyrLysArg 467
1884 GACCAATGACAGCAACAAAGACAGATATATGAACTTAATGAAAA 1933
467 gGluIleTyrGlyArgGlyLeuProGluArgGlnLysGlyGlnLeuAla 484
1934 AGAATATATGTCGAGGCTCTCCAGAAAGCAAAAGGCGAGCTACCTG 1983
484 aGlnArgAlaLysGlnValGluGluPheLeuGlnArgLysArgGluAla 500
1984 TAGAAAGAGCTAAACAAAGTAGAAGATTCTGACGCGAAAGGGAGCT 2033
501 MetGlnAsnLysAlaArgAlaGlnGluLysHisMet..... 511
2034 ATGCAATATTAAGCTGACGCCAGACCATATNGGAATTCGCAAAAC 2083
511 ..... 511
2084 GCGAGCTATGTATGAGCGCAGGCCAGCTTTCAAGAGAGAGAGCA 2133
512 .....ValTyrLeuAlaArgLeuArgIleArgLeu 522
2134 GAAACAAAGAGAGAGGTTTATCTGCAAGACTGAGCAAAATTAAGCTA 2183
523 GlnAsnPheAsnGluArgGlnGlnIleLysAlaLysLeuArgGlyGly 539
2184 CAGAAATTTCAATGAGCGCCCAACAGATTAAAGCAAACTTCGAGTAA 2233

```

539 sLYSGluAlaasnHisSerGluGluGluGluGluSerGluGluAlaAspK 556
 |||||
 2234 GAAGAACAACCTAATCATCTCTGAAGACAAAGAAAGAAAGAGAGCTGAC 2283
 556 eTArgArg.LysLysIleGluSerLeuLysAlaHisAlaasnAlaArgAl 572
 |||||
 2284 TGAGCGCAAAAAAATGATCACTGAGGCCCATGCATCAATGCACGTGC 2333
 572 aAlaValLeuLysGluGluGluGluGluGluGluGluGluGluGluGlu 589
 |||||
 2334 TCCTCTACTAAAGACAACTAGAACCAAGAAAGAAAGAGCTTATGACA 2383
 589 rGluLysValLeuTrpGluGluHisLeuValAlaLysGlyValLysSer 605
 |||||
 2384 GGAAGAAAGAGTGGAGAGCAATGGGCTTAAGAGATTAAAGAGT 2433
 606 SerAspValSerProProLeuGluHisGluThrGlyLysSerProse 622
 |||||
 2434 TCTGATGTTTCTCCACTTTGGACAGCATGAAACAGTGGCTCCATC 2483
 622 rLYSGluGlnMetArgSerValIleSerValThrSerAlaLeuLysGluY 639
 |||||
 2484 AAAGCAACAGATGAGATCTGTTATTTCTGTAACCTTCAGCTTGAAGAAG 2533
 639 aIGlyValAspSerSerLeuThrAspThrArgGluThrSerGluGluMet 655
 |||||
 2534 TTGGCGTGACAGTAGTTTAAGTACCGGAACTTCAGAAAGAGATG 2583
 656 GlnLysThrAsnAsnAlaIleSerSerLysArgGluIleLeuArgArgLe 672
 |||||
 2584 CAAAAGACCAACAAATGCTATTTCAAAGTAAAGCAAAATACTTGGCAGATT 2633
 672 uASnGluAsnLeuLysAlaGlnGluAspGluLysGlyMetGlnAsnLeuS 689
 |||||
 2634 AAATGAAATCTTAAAGCTCAAGAAAGTAAAGAAAGAAATCAGAAATCTCT 2683
 689 eAspThrPheGluIleAsnValHisGluAspAlaLysGluHisGluLys 705
 |||||
 2684 CTGATACCTTTGAGATAAATGTCATGAAGATGCCAAAGACATGAAAAA 2733
 706 GlnLysSerValSerSerAspArgLysLysTrpGluAlaGlyGluLe 722
 |||||
 2734 GAAAAATAGATTTCATCTGATCCAGAAAGTGGAGGAGAGGAGTCAACT 2783
 722 uValIleProLeuAspGluLeuThrLeuAspThrSerPheSerThrThrg 739
 |||||
 2784 TGTGATTCCTCGATGAGTTAACACTAGAAATCTCTCTCAACAACCTG 2833
 739 LuArgHisThrValGlyGluValIleLysLeuGluProAsnGlySerPro 755
 |||||
 2834 AAAGACATACACAGTGGAGAACTTATTAATTAAGTCTTAATGCATGCCA 2883
 756 ArgArgAlaIleTrpGlyLysSerProThrAspSerValLeuLysIleLeuG 772
 |||||
 2884 AGAAGAGCGCTGGGAAAAAGTCCACAGATTCCTCTTAAGAAATCTGG 2933
 772 yGluAlaGluLeuGluLeuGlnThrGluLeuGluAsnThrThrIleA 789
 |||||
 2934 AGAACTGAACTACAACTTCAGACAGAACTATTGAAAAAATACAACTATTA 2983
 789 rGSerGluIleSerProGluGlyGluLysTrpLysProLeuIleThnGly 805
 |||||
 2984 GAAGTGAGATTTCCTCCAGAGGGGAAAAAGTACAAACCTTAATTAAGTGA 3033
 806 GlnLysLysValGlnCysIleSerHisGluIleAsnProSerAlaIleVa 822
 |||||
 3034 GAAAAAAGATACATGATTTCACATGAAATTAACCAATCAGCTATTGT 3083
 822 LAspSerProValGluThrLysSerProGluPhSerSerGluAlaSerProG 839
 |||||
 3084 TGAATTCCTCCTGTTGAGCAAAAAAGTCCGAGTTCAAGTGAAGCAATCTCAC 3133

839 lMetSerLeuLysLeuGluGlyAsnLeuGluGluProAspAspLeuGlu 855
 |||||
 3134 AGATGCTATTGAAACCTGGAAGAAATTTAGAAAGAACCTGATGATTGGAA 3183
 856 ThrGluIleLeuGluGluProSerGlyThrAsnLysAspGluSerLeuPr 872
 |||||
 3184 ACAGAAATTTCTACAGAGCCAAAGTGAACAAACAAAGATGAGAGCTTCC 3233
 872 oCysThrIleThrAspValTrpIleSerGluGluLysGluThrLysGluT 889
 |||||
 3234 ATGCACTATTACTGATCTGTGGATTGTGAGGAAAAAAGAAACAAAGAAA 3283
 889 hrcGlnSerAlaAspArgGlyIleThrIleGlnGluAsnGluValSerGluSp 905
 |||||
 3284 CTCAGTCGCGAGATAGGATCCCATTCACGAAAGAAAGAACTTCTGAAGAT 3333
 906 GlyValSerSerThrValAspGlnLeuSerAspIleHisIleGluProG 922
 |||||
 3334 GGAATCTCGAGTACTGTGACCAACTTAAGTATGATCATATAGAGCCCTGG 3383
 922 yThrAsnAspSerGlnHisSerLysCysAspValAspLysSerValGlnP 939
 |||||
 3384 AACCAATGATTCTCAGCACACTTAAATGATGATGATGATGATGATGATG 3433
 939 rGluLupProPhePheHisLysValValHisSerGluHisLeuAsnLeuVal 955
 |||||
 3434 CGGAACCATTTTCCATAAAGTGTTCATTCTGAACACTGAACTTAAGTTC 3483
 956 ProGlnValGlnSerValGlnCysSerProGluGluLysSerPheAlaPheAr 972
 |||||
 3484 CCTCAATTCATACATGATGCTGATGCTTCCACCAAGAAATCCTTGCATTTCG 3533
 972 gSerHisSerHisLeuProProLysAsnLysAsnLysAsnSerLeuLeuI 989
 |||||
 3534 ATCTCACTGCATTTACCCCAAAAAATTAACAAAGATTTCTCTTGCGA 3583
 989 lGlyLeuSerThrGlyLeuPheAspAlaAsnAsnProLysMetLeuArg 1005
 |||||
 3584 TTGGACCTTTCACTGCTGCTTTGATGACAAACCAACCAAGATGTTAAGG 3633
 1006 ThrCysSerLeuProAspLeuSerLysLeuPheArgThrLeuMetAspVa 1022
 |||||
 3634 ACATGTTCACTTCACATCTCTCAAGCTGTTCAAGAACCTTATGATGAGT 3683
 1022 lProThrValGlyAspValArgGlnAspAsnLeuGluIleAspGluIleL 1039
 |||||
 3684 TCCACCGGTAGGAGATGTTGTCAGACAAATCTTGAATAATGATGAATTTA 3733
 1039 yAspAspGluAsnIleLysGluGlyProSerAspSerGluAspIleValPhe 1055
 |||||
 3734 AAGATGAAACATTAAGAAAGACCTTCGATTCGAAACACTTGTGTTT 3783
 1056 GlnGluThrAspThrAspLeuGlnGluGlnAlaSerMetGlnGluLe 1072
 |||||
 3784 GAAAGAACTGACACAGATTTACAAAGCTGCAAGGCTCATGTAACAGATT 3833
 1072 uLeuArgGluGluProGlyGluGluTrpSerGluGluGluGluSerValL 1089
 |||||
 3834 ACTTAGGGAACAACCTGCTGAAGAAATACAGTGAAGAAAGAGTGAAGTCT 3883
 1089 eLysAsnSerAspValGluProThrAlaAsnGlyThrAspValAlaAsp 1105
 |||||
 3884 TGAAGAACAGTATGATGAGCCCAACTGCAAAATGGGACAGATGTGCAAGAT 3933
 1106 GlnAspAspAsnProSerSerGluSerAlaLeuAsnGluGluTrpHisSe 1122
 |||||
 3934 GAAATGACAAATCCCACTAGTGAAGAGTCCCTGAAGAAAGAAATGACACTC 3983
 1122 rAspAsnSerAspGlyGluIleAlaSerGluCysGluCysAspSerValP 1139
 |||||
 3984 AGATTAACAGTATGAGTGAATTTCTAGTGAATGTGAATCTCGATAGTGTCT 4033
 1139 hAsnHisLeuGluGluLeuArgLeuHisLeuGluGluGluMetGlyPhe 1155

```
|||||
4034 TTTAACCTTTTGGAGAACTGAGACTTCATCTGAGAGAAATGGGCTT 4083
1156 GtUysPhepHeGluValTyrgUlysIleUysAlaIleHisGluAspG 1172
|||||
4084 GAAAAATCTTGTGAGGTTATGAGAAATAAAGCTATTTCATGAAAGATGA 4133
1172 uAspGluAsnIleGluIleCysSerLysIleValGlnAsnIleLeuGlyA 1189
|||||
4134 AGATGAAATATTGAAATTTGTTCAAAATAGTCAAAATATTTTGGGAA 4183
1189 snGluHisGlnHisIleUysTyrgAlaLysIleLeuHisIleUysValMetAlaAsp 1205
|||||
4184 ATGAACTACAGCACTTATGCAAGATCTTCATTCATTACTGTCAGAT 4233
1206 GtYalAtyrgInGluAspAsnAspGlu 1214
|||||
4234 GGAGCTTACCAAGAAATATGATGAA 4260

seq_name: gb_pat:AX224473

seq_documentation_block:
LOCUS AX224473 3024 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 5 from Patent WO0161016.
ACCESSION AX224473
VERSION AX224473.1 GI:15554707
KEYWORDS
SOURCE
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 3024)
Wolke,D.W., Hu,Y., Nepomichy,B., Turner,C.A. and Zambrowicz,B.
Human kinases and polynucleotides encoding the same
Patent: WO 0161016-A 5 23-AUG-2001;
JOURNAL
Lexicon Genetics Incorporated (US)
FEATURES
location/Qualifiers
1..3024
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1115 a 503 c 681 g 725 t
ORIGIN

alignment_scores:
Quality: 5162.00 Length: 1007
Ratio: 5.126 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-783-320-4 x AX224473 ..

Align seg 1/1 to: AX224473 from: 1 to: 3024

208 MetLysAsnLeuValLeuLysIleIleSerGlySerPheProValSe 224
1 ATGAAAAACCTGTACTGAAATATATCTGATCTTTCCACCTGATGTC 50
224 rLeuHisTySerTyAspLeuArgSerLeuValSerGlnLeuPheLysA 241
|||||
51 TTTCATATTATCTATGATCTCCGAGATTGGTCTCAGTTATTAAAA 100
241 rGAsnProArgAspArgProSerValAsnSerIleLeuGluLysGlyPhe 257
|||||
101 GAAATCTTAGGATGAGCATGACATCACTCATTCATTTGAGAAAAGGTTT 150
258 rLeuLysArgIleGluLysPheLeuSerProGlnLeuIleAlaGluG 274
|||||
151 ATACCAAAAGCATTTGAAAAGTTTCTCTCTCCTCAGCTTAATGCAGAGA 200
274 uPheCysLeuLysThrPheSerTyPheGlySerGlnProIleProAlaL 291
|||||
201 ATTTGTCATAAAACATTTTGGAGATTGATCACAGCCTAATACCAGCTA 250

291 ySarGProAlaSerGlyGlnAsnSerIleSerValMetProAlaGlnLys 307
|||||
251 AAATACCAAGCTTCAGGACAAACATCGATTCTGTATGCTGCTCAGAAA 300
308 rLeuThrLysProAlaAlaLysTyrgLysIleProLeuAlaTyrgLysTy 324
|||||
301 ATACCAAGCTTCGCGCTTAATGGAATACCTTTAGCATATTAAGAAATA 350
324 rGlyAspLysLysLeuHisGlnLysTyrgProLeuGlnLysHisLysGlnA 341
|||||
351 TGGAGATTAATAATTTACAGAAAGAACCACTCGCAAAACATAAACAGG 400
341 rAlaHisGlnThrProGluLysArgValAsnThrGlyGluGluLysArgProLys 357
|||||
401 CCCATCAAAATCCAGAAAGAGAGTGAATACGTGAGAAAGAAAGAGGAAA 450
358 rLeuSerGluGlnAlaAlaArgLysArgArgLeuPheIleGluLysG 374
|||||
451 ATATCTGAGGAAGCAGCAAGAAAGAGAGCTGAAATTTATGAAAGAAA 500
374 uLysLysGlnLysAspGlnIleIleSerLeuMetLysAlaGlnMetL 391
|||||
501 AAAGAAACAAAAGGATCAGATTATTAGTTAATGAAGGCTGAACAAATGA 550
391 ySarGInGluLysGluArgLeuGluArgIleAsnArgAlaArgGluGln 407
|||||
551 AAAGCAAGAAAGAAAGAGGTTGGAAAGAAATTAATAGGCGCAGGAAACA 600
408 GtYTrpArgAsnValLeuSerAlaGlyIleSerGlyGluValLysAlaPr 424
|||||
601 GGATGCAAGAAATGCTCTAAGCGTGGTGAAGGCTGAATGAATGAAGCTCC 650
631 yArgGlnGluLysGluArgLeuGluArgIleAsnArgAlaArgGluGln 441
|||||
441 oPheLeuGlySerGlyThrIleAlaProSerSerPheSerSerArgG 457
|||||
651 TTTTCTGGCAGTGCAGGAGACTATAGCTCATCTTTCTTCTGCGAG 700
441 yGlnIleTyrgLinhIstYrHisAlaIlePheAspGlnMetGlnGlnArg 474
|||||
701 GACAGTATGACATTTACCTGATGCTATTTTGCAAAAGCAGCAACAAAAG 750
458 rAlaGluAspAsnGluAlaLysTrpLysArgGluLysArgGlyLe 474
|||||
751 GCAGGAAGATATGAACTAATGAAAGAAAGAAATATATGTCGAGGTC 800
474 uProGluArgGlnLysGlyGlnLeuAlaValGluArgAlaLysGlnValG 491
|||||
801 TCCAGAAAGGCNAAGGCGACCTAGCTGTAGAAAGCTTAACCAAGTAG 850
491 rGluLysPheLeuGlnArgLysArgGluAlaMetGlnAsnLysAlaArgAla 507
|||||
851 AAGGTTCTCGCAGCAAAACGGAAGCTATGCAAAATTAAGCTCGAGCC 900
508 GtUglYHisMetValTyrgLeuAlaArgLeuArgGlnIleArgLeuGlnAs 524
|||||
901 GAAGGACATATGTTTATCTGCAAGACAGAGCAAAATTAAGACTACAGAA 950
524 nPheAsnGluArgGlnGlnIleLysAlaLysLeuArgLysLysLysG 541
|||||
951 TTTCATAGACGCCCAACAGATTAAAGCCCAACTGCTGCTGATAAAGAAAG 1000
541 rAlaIleAsnHisSerGluGluGlnGlnLysSerGluGlnAlaAspMetArg 557
|||||
1001 AAGCTAATCATCTTGAAGGACCAAGAAAGAAAGAAAGAGCTGACATGAG 1050
558 rArgLysLysIleGluSerLeuLysAlaHisAlaAsnAlaArgAlaAla 574
|||||
1051 CGCAAAAAGAAACGAACTACGAAAGGCCCATGCAAAATGCACGCTGCTGT 1100
574 rLeuLysGluGlnLeuGluArgLysArgLysGluLysAlaTyrgGluArgGlu 591
|||||
1101 ACTAAAAGAAACACTTAGAAGAAAGAAAGAGGCTTATGAGAGAGAAA 1150
```

```
591 yslvsvaltrpglugluhslseuvalalalsgylvallysserserasp 607
|||||
1151 AAAAGGTGGGAGAGCATTTGGCTGAAGAGATTAAGAGTTCGAT 1200
608 valserproleuglycnihsigluhtrgclgyserserserasp 624
|||||
1201 GTTTCACACCTTTGGACACATGAAAGAGTGGCTCTCCATCAAGCA 1250
624 nglmetargservalleservalthrseralaleuylglvalgly 641
|||||
1251 ACAGATGAGATCTGTTATCTCTGACTACGCTTGAAGAAAGTGGCG 1300
641 alaserseuvalthraspthrargluhtrsergluglumetglnlys 657
|||||
1301 TGGACAGATGTTTAACTGATACCGGGAACCTTCAGAAAGATGCAGAA 1350
658 thrasnasnalaileserserlyargluhtrleuargargleuasn 674
|||||
1351 ACCAACAATGCTATTTCAGTAAAGCAAGAAATCTTCAGATTTAAATGA 1400
674 usasnleuylsalaglugluaspglulysgylmetglnasnleuaser 691
|||||
1401 AATCTTAAAGCTCAAGAAATGAAAAAGAAATGCAGAAATCTCTGATA 1450
691 hrphgcluleasnvalhsigluaspalalysgluhslgylulys 707
|||||
1451 CTTTGAATTAATCTTCATGACATGCCAAAGCATGAAAAAGAAAA 1500
708 servalseraspargllystrpdlualaglyglnleuval 724
|||||
1501 TCGAGTTTCATCTGATCGCAAGAGTGGAGAGAGTCACTGATGAT 1550
724 eproleuaspgluleuvalthrleuaspthrserphsethrthrcluar 741
|||||
1551 TCCCTGGAGTGAATACATGATACATCTCTCTACAACTGAAGAAC 1600
741 lsfthrvalglgluvalilleuylseuuglyproasnlyserproarg 757
|||||
1601 ATACAGTGGAGAAATTAATAATTAAGTCTTAATGATCTCCAAAGAA 1650
758 alatrpglyllyserprothraspservalleuylseuuglylual 774
|||||
1651 GCCTGGGGGAAAAAGTCCGACAGATTCGTCTTAAAGATCTTGAGAG 1700
774 agluleuugluuugluhtrglnleuugluasnthrthrleargserg 791
|||||
1701 TGAATCAACAATTCCAGACAGAACTATTGAATAACAATATTAGAGTG 1750
791 luileserprogluglylulysyllysproleuilethrnglylulys 807
|||||
1751 AGATTTCCTCCGAAAGGAAAAAGTCAAAACCTTAATTACTGGAGAAAA 1800
808 lysvalglnlylleserhisglnleasnproseralalevalasp 824
|||||
1801 AAAGTACAAATGATTTCATGATGAATAACCACTGATGATGTTGATTC 1850
824 rprovalgluthrlyserprogluphserglualaserproglumets 841
|||||
1851 TCCCTGTGAGCAAAAAAGTCCGAGTTCAGTGAAGCATCTCCACAGATG 1900
841 erleuylseuugluuylasnleuugluuuproaspaspserleuuglu 857
|||||
1901 CATTGAACCTGGAAAGAAATTTAGAAAGAACTGATGATTTGAAAGCAA 1950
858 lleuuglnlugluprosergllythrlyasnlyspgluserserleu 874
|||||
1951 ATTCTACAAAGAGCCAAAGTGAACAAACAAAGATGAGAGCTTGCCATG 2000
874 rlethrspvaltrpilesersglnugluhtrlysgluhtrglns 891
|||||
2001 TATTACTGATGTGTGATTAAGAGAAAAAGAAACAAAGAAACTAGT 2050
891 erhlasparglilethrilleuglnuasngluvalsergluaspglyal 907
|||||

|||||
2051 CGGACAGATAGATCACCATTGACAGAAATGAAGTTCTGAAATGAGAGTC 2100
908 sersethrvalaspgluleuaseraspillehsllegluuproglythr 924
|||||
2101 TCGAGTACTGTGACCACTTATGACATTCATATGAGCTCGAAGCA 2150
924 naspsersglnhisserlyscysaspyalaspyservalglnproglup 941
|||||
2151 TGATTCACACCTTAATGATGATGATGATGATGATGATGATGATGAT 2200
941 rophehsllyslvalalhissergluhslseuvalproglin 957
|||||
2201 CATTTCATTAAGTGGTGTTCATTCGAACTGAACTGAACTGATCCCTCA 2250
958 valglnsevalglnlyserproglugluserphealapheargserr 974
|||||
2251 GTTCAATCAGTTCAGTGTCCACAGAAAGATCTTCGATTCGATTCGA 2300
974 sserhisleuuproprolysasnlysnlysnlysnlysnlysnlysn 991
|||||
2301 CTCCCATTTACACCAAAAAATAAAAAACAAGATTCCTTGATGGAC 2350
991 euserthrnglyleupheaspalasnasnprolysmetleuargthrc 1007
|||||
2351 TTTCACTGGTCTGTTGATGCCAAACACCAAGATGTTAAGGACATGT 2400
1008 serleuproaspserlyseuvalpheargthrleuvalaspvalproth 1024
|||||
2401 TCACCTTCAGATCTCTAAGCTGTTCCAGACCTTATGATGATGCCAC 2450
1024 rvalglyaspvalarglnaspasnleuugluilleaspgluileylasp 1041
|||||
2451 CGTGGAGATGTTGTTAAGCAATCTGTAAGATGATGTAATTAAGATG 2500
1041 luasnilleuglyllyproseraspsergluaspllevalphecluglu 1057
|||||
2501 AAAACATTAAAGAAAGACCTTCGATTCGAAACCTTCGTTTGAAGAA 2550
1058 thraspthraspleuuglnleuuglnalasermetglnleuvalthr 1074
|||||
2551 ACTGACACAGATTTTAAAGAGTGCAGGCTCGATGGAACAGTACTAG 2600
1074 ggluglnproglugluuyltrserglugluglugluservalleuylsa 1091
|||||
2601 GGAACAACCTGGTGAAGATATACAGTGAAGAAAGAGTCACTGTTGA 2650
1091 snseraspvalgluprothralsasnlythrlyaspvalalaspgluas 1107
|||||
2651 ACAGTATGTGAGCCCAACTGCMAATGGGACAGATGTGGCAGATGAAT 2700
1108 aspsasnprosersergluseralaleuasnugluhtrphiserser 1124
|||||
2701 GACATGCCAGTATGAAAGTCCCTGACAGAAAGATGSCATCAGATTA 2750
1124 nseraspglugluilleaserclyscylucysaspservalpheasn 1141
|||||
2751 CAGTATGAGTGAATTTCTAGTAAATGTGATGATGATGATGATGAT 2800
1141 lisleuuglnleuargluuylseuuglnleuuglnleuuglylphelg 1157
|||||
2801 ATTTAAGAGAACTGAGACTTCATCTGAGAGAGGAAATGGGCTTTGA 2850
1158 phehegluvaltyrglylulysllelysalallehslgluasgluas 1174
|||||
2851 TTTCTTGGAGTTTATGAGAAATTAAGGCTATTCATGAGATGAGATGA 2900
1174 uasnilleugluilecyserserlylevalglnasnilleuuglyuasng 1191
|||||
2901 AAATATTGAATTTGTTCAAAAAATAGTCAAAATATTTTGGAAATGAAC 2950
1191 lsglnhisleuvalalylsilleuunhisleuvalmetlasparglyal 1207
|||||
```

```

2951 ATCAGACTTTATGCCAGATTCCTTCATTAGTCATGCGAGATGAGCC 3000
1208 TGTGInGUAspAsnAspGU 1214
|||||
3001 TACCAGAAAGATTAATGATGAA 3021

seq_name: gb_pat:AX224469

seq documentation block:
LOCUS AX224469 3108 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 1 from Patent WO0161016.
ACCESSION AX224469
VERSION AX224469.1 GI:15554705
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 3108)
AUTHORS Walke,D.W., Hu,Y., Nepomitchy,B., Turner,C.A. and Zambrowicz,B.
TITLE Human kinases and polynucleotides encoding the same
JOURNAL Patent: WO 0161016-A 1 23-AUG-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source
1. 3108
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1145 a 520 c 708 g 735 t
ORIGIN

alignment_scores:
Quality: 5138.00 Length: 1035
Ratio: 5.102 Gaps: 1
Percent Similarity: 97.295 Percent Identity: 97.295

alignment block:
US-09-783-320-4 x AX224469 ..

Align seg 1/1 to: AX224469 from: 1 to: 3108

208 MetLysAsnLeuValIleuLysIleIleSerGlySerPheProProValSe 224
|||||
1 ATGAAACACCTGGTACTGAGATTAATCTCGATCTTTTCCACCTGTGTC 50
|||||
224 rLeuHisTyrSerTyrAspLeuArgSerLeuValSerGlnLeuPheLysA 241
|||||
51 TTGGATTATTCCTATGATCTCCGAGTTGGTGTCTCACTTATTTAAAA 100
|||||
241 rGAsnProArGAspArgProSerValAsnSerIleLeuGluLysGlyPhe 257
|||||
101 GAAATCTCAGAGATGACCACTCACTCATTTGGAGAAAGGTTT 150
|||||
258 IleAlaLysArgIleGluLysPheLeuSerProGlnLeuIleAgluG 274
|||||
151 ATAGCCAAACGCATTGAAAGTTCTCTCTCTCAGCTTATTCAGAAAGA 200
|||||
274 uPheCysLeuLysThrPheSerLysPheGlySerGlnProIleProAla 291
|||||
201 ATTTGTCTTAAACATTTTCGAACTTGGATCAGCTATATGCACACTA 250
|||||
291 ySarArgProAlaSerGlyGlnAsnSerIleSerValMetProAlaGlnLys 307
|||||
251 AAAGACCAGCTTCAGACAAACCTCGATTCTGTTATGCCGCTCAAGAAA 300
|||||
308 IleThrLysProAlaLysTyrGlyIleProLeuAlaTyrLysLysTyr 324
|||||
301 ATTAACAAGGCTGCCGCTAAATATGAAATACCTTTAGCATATTAAGAAATA 350
|||||
324 rGlyAspLysLysLeuHisGluLysLysProLeuGlnLysHisLysGlnA 341
|||||
351 TGGAGATTAATAAATTACAGAAAGAAACCACTGCAGAAAACATTAACGCG 400

```

```

341 IaHisGlnThrProGluLysArgValAsnThrGluGluGluArgArgLys 357
|||||
401 CCCATCAAACTCCAGAGAGAGTGAATATCTGAGAGAAAGAGAGAAA 450
|||||
358 IleSerGluGlnLysAlaLysArgArgArgLeuGluPheIleGluLysG 374
|||||
451 ATATCTGAGAGACAGCAGAAAGAGAGAGAGCTGGAATTTATTTGAAAAA 500
|||||
374 uLysLysGlnLysAspGlnIleIleSerLeuMetLysAlaGlnMetL 391
|||||
501 AAAAGAAACAAAGAGATCAGATTATTAAGTTAGAAAGGTGACCAATGA 550
|||||
391 ySarArgGlnLysGluArgLeuGluArgIleAsnArgAlaArgGluGln 407
|||||
551 AAAGGCAAGAAAGAAAGAGGTTGAAAGAAATAAATAGGCGGAGACAA 600
|||||
408 GlyTPArgAsnValLeuSerAlaGlySerGlyGluValLysAlaP 424
|||||
601 GGATGGAGAAATGTGTAAGTGTGGTGGAGAGTGTGAAGTAAAGGCTCC 650
|||||
424 oPheLeuGlySerGlyGlyThrIleAlaProSerSerPheSerSerArg 441
|||||
651 TTTTCTGGCAGTGGAGGACATATAGCTCATCATCTTTTCTCTCGAG 700
|||||
441 yGlnTyrGluHisTyrHisAlaIlePheAspGlnMetGlnGlnArg 457
|||||
701 GACAGTATGAACATTACCATGCCATTTTGTGACCAATGCAGCAAGAAAGA 750
|||||
458 IAAGluAspAsnGluAlaLysTyrPlyArgGluIleTyrGlyArgGlyLe 474
|||||
751 GCGAAAGATTAATGAAGCTAAATGAAAGAAAGAAATATATGTCGAGGCT 800
|||||
474 uProGluArgGlnLysGlyGlnLeuAlaValGluArgAlaLysGlnValG 491
|||||
801 TCCAGAAAGGCCAAAGAGCGCAGCTGATGAAAGAGCTTAAACAAGTAG 850
|||||
491 IuGluPheLeuGlnArgLysArgGluLysMetGlnAsnLysAlaArgAla 507
|||||
851 AAGAGTTCTTCGAGCAAGAAAGGGAAGACTATGCAGATTAAGCTCGAGCC 900
|||||
508 GlnGlyHisMet..... 511
|||||
901 GAAGGACATATGGAATCTCGAAACCTGGCAGCTATGTATGAGAGCAG 950
|||||
512 .....ValT 513
|||||
951 GCCCAGCTTTCMAAGAGAGGAAAGCCAGAAACAAAGAGAGAGGTTT 1000
|||||
513 yLeuAlaArgLeuArgGlnIleArgLeuGlnAsnPheAsnGluArgGln 529
|||||
1001 ATCTGGCAGACTGAGGCAAAATAGACTACAGAAATTCATATGAGCGCCA 1050
|||||
530 GlnIleLysAlaLysLeuArgGlyGluLysLysGlnLysAlaAsnHisSerG 546
|||||
1051 CAGATTAAAGCCAACTCTGCTGTGTAAGAAAGAAAGAACTATATCTCTGA 1100
|||||
546 uGlyGlnGluGlySerGlnGluAlaAspMetArgArgLysLysIleGluLys 563
|||||
1101 AGGACAAAGAAAGAAAGTGAAGAGGCTGACATGAGCGCAAAAAAATCGAAT 1150
|||||
563 eLeuLysAlaHisAlaAsnAlaArgAlaValIleLysGluGlnLeu 579
|||||
1151 CACTGAAGGCCCATGCAATGACAGCTGCTGCTGACTTAAAGAGACACTA 1200
|||||
580 GlnArgLysArgLysGlnLysTyrGluArgGluLysLysValITrPgluG 596
|||||
1201 GAACGAAAGAGAAAGAGGCTTATGAGAGAGAGAAAAAGCTGGAGAGA 1250
|||||
596 uHisLeuValAlaLysGlyValLysSerSerAspValSerProProLeu 613
|||||
1251 GCATTGTGGTCTTAAAGAGATTAAAGGTTCATATGTTCTCCACCTTTGG 1300
|||||
613 yGlnHisGlnThrGlyGlySerProSerLysGlnGlnMetArgSerVal 629

```

```

|||||
1301 GACACACGAGAAACAGTGGCTCTCCATCAAGCAACAGATGAGATCTGTT 1350
630 IIESErValThrSerAlaLeuLySGluValGlyValAspSerSerLeuTh 646
1351 ATTCTGTACCTTCAGCTTTGAAGAAGTTGGCGTGGACAGTGGTTAAC 1400
646 rASpThrArgLysThrSerGluGluMetGluLysThrAsnAlaIleS 663
1401 TGATACCCGGGAACCTTCAGAAAGATGCCAAAGACCAACAATGCTATT 1450
663 eSErLysArgGluIleLeuArgArgLeuAsnGluAsnLeuLysAlaGlu 679
1451 CAAGTAAGCGAAGAACTCTCGAGATTAAATGAATCTTAAAGCTCAA 1500
680 GluAspGluLysGlyMetGluAsnLeuSerAspThrPheGluIleAsnVa 696
1501 GAAGATGAAAAAGAAATGCAGAAATCTCTGATACTTTGAGATTAATGT 1550
696 IHisGluAspAlaLysGluHisGluLysGluLysSerValSerSerAspa 713
1551 TCATGAAGATGCCAAAGAGCATGAAAAAAGAAATCAGTTTCATCTGATC 1600
713 rGLysLysTrpGluAlaGlyGluLeuValIleProLeuAspGluLeu 729
1601 GCAGAAAGTCGGAGCGACAGAGTCACATGATTCCTCTGATGAGTTA 1650
730 ThrLeuAspThrSerPheSerThrThrGluArgHisThrValGlyGluVa 746
1651 ACACTAAGATACATCTCTCTACAACTGAAAGACATACACTGGAGAGAGT 1700
746 ILeuLysLeuGlyProAsnGlySerProArgArgAlaTrpGlyLysSer 763
1701 TATTAAATTAGTCTCTAATGATCTCCAGAAAGCCTGGGGGAAAAAGTC 1750
763 rOTHrAspSerValLeuLysIleLeuGlyGluAlaGluLeuGlu 779
1751 GCACGATCTGTCTAAGATACCTGAGAGAGCTGAACATCAACATCTCAG 1800
780 ThrGluLeuGluAsnThrThrIleArgSerGluIleSerProGlu 796
1801 ACAGACATATTAGAAATATACAACTATTAGAGAGATTTCTCCCGAGG 1850
796 yGluLysTrpLysProLeuIleThrGlyLysLysValGlnCysIleS 813
1851 GGAAGAAAGTAAACCCCTTAATTACTGAGAAAAAAGTACAATGTATT 1900
813 eHHisGluIleAsnProSerAlaIleValAspSerProValGluThrLys 829
1901 CACATGAATATAACCCATCAGCTATTGTGATTCCTCTGTGAGACAAA 1950
830 SerProGluLysPheSerGluAlaSerProGluMetSerLeuLysLeuGlu 846
1951 ACTCCCGAGTTCAGTGAAGCATCTCCACAGATGTCATTGAAACTGGAGG 2000
846 yAsnLeuGluGluProAspAspLeuGluThrGluIleLeuGluGluPro 863
2001 AATATTAGAAACCTGATGATTGGAAACAGAAATTTCTAAAGAGCCAA 2050
863 eGlyThrAsnLysAspGluSerLeuProCysThrIleThrAspValTrp 879
2051 GTGGACAACAAGATGAGAGCTTCCATGCTACTATTACTGATGTGTGG 2100
880 ILeSerGluGluLysGluThrLysGluThrGlnSerAlaAspArgIleTh 896
2101 ATTACTGAGGAAAAAGAAAGAAAGAACTCAGTCGGCAGATAGGAGTAC 2150
896 rIleGlnGluAsnGluValSerGluAspGlyValSerSerThrValAspG 913
2151 CATTCAGGAAATGAAGTTTCTGAAGATGAGAGTCTCGAGTACCTGTGACC 2200
913 ILeuSerAspIleHisIleGluProGluThrAsnAspSerGlnHisSer 929
|||||

```

```

2201 AACTAGTCACATTTCATATAGAGCCCTGAAACCAATGATTCAGACACTCT 2250
930 LysCysAspValAspLysSerValGlnProGluProPhePheHisLysVa 946
2251 AAATGTGATGATAGTAAGTCTGTGCAACCGGAACCATTTTCTCAAAAGT 2300
946 yValHisSerGlnHisLeuAsnLeuValProGlnValGlnSerValGlnC 963
2301 GGTTCATTCTGAACTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCT 2350
963 ySerProGluGluSerPheAlaPheArgSerHisSerHisLeuProPro 979
2351 GTTCACCAAGAAATCTTTGCAATTCGATCTCCTCCATTCGATTCACCA 2400
980 LysAsnLysAsnLysAsnSerLeuLeuIleGlyLeuSerThrGlyLeuPh 996
2401 AAAAATAAAAACAAGATTCCTGTGATGGACCTTCAACTGCTGTCT 2450
996 eAspAlaAsnAsnProLysMetLeuArgThrCysSerLeuProAspLeuS 1013
2451 TGATGCCAAACCAACCAAGATGTTAAGACATGTTCCACTTCAGATCTCT 2500
1013 eTrpLysLeuPheArgThrLeuMetAspValProThrValGlyAspValArg 1029
2501 CAAAGCTGTTCAAGAACCTTATGATGATGCCACCGTAGAGATGTTGCT 2550
1030 GluAspAsnLeuGluIleAspGluIleLysAspGluAsnIleLysGlu 1046
2551 CAAGACAAATCTTGAAATAGTGAATTAAGATGAAGAAACATTAAGAGAG 2600
1046 yProSerAspSerGluAspIleValPheGluGluThrAspThrAspLeuG 1063
2601 ACCTTCGATCTGAAAGACATTTGTTGAAAGAACTGACACAGATTTAC 2650
1063 IArgLysLeuGlnAlaSerMetGluGluLeuLeuArgGluGluProGly 1079
2651 AAGAGCTGAGCGCTCGATGAGACAGTACTTAAAGAACACCTGGTGA 2700
1080 GluTrpSerGluGluGluGluSerValLeuLysAsnSerAspValGlu 1096
2701 GAATACAGTGAAGAAAGAGTCACTTGAAGAACAGTATGTGGACC 2750
1096 oThrAlaAsnGlyThrAspValAlaAspGluAspAsnProSerSerG 1113
2751 AACTGCAAAATGGGACAGATGTGCGAGATGAGATGCAATCCACATAGTG 2800
1113 IuSerAlaLeuAsnGluGluTrpHisSerAspAsnSerAspGlyLysIle 1129
2801 AAATGCCCTGAAAGAAATGCGCACTCAGATTAACAGTATGATGTAAT 2850
1130 AlAserGluCysGluCysAspSerValPheAsnHisLeuGluLeuArg 1146
2851 GCTAGCAATGTGAATGCGATGCTTAAACATTTAGAGAACTGAG 2900
1146 gLeuHisLeuGluGluGluMetGlyPheGluLysPhePheGluValArg 1163
2901 ACTTCATCTGAGACAGAAATGGCTTTGAAAAATCTTTGAGGTTTATG 2950
1163 IuLysIleLysAlaIleHisGluAspGluAspGluAsnIleGluIleCys 1179
2951 AGAAATAAAGGCTATTCAAGATGAAGATGAAGATTAATTAATTTG 3000
1180 SerLysIleValGlnAsnIleLeuGlyAsnGluHisGlnHisLeuTrpAl 1196
3001 TCAAAAATAGTTCAAAAATATTTGGGAAATGAACATCACACATCTTATG 3050
1196 aLysIleLeuHisLeuValMetAlaAspGluAlaTrpGlnGluAspAsn 1213
3051 CAAGATCTTCATTTAGTATGATGACAGATGAGAGCCTAACAAAGATTA 3100
1213 spGlu 1214
3101 ATGAA 3105

```



```

seq_name: gb_ro:S45828
seq_documentation_block:
LOCUS      S45828                4263 bp    mRNA           linear    ROD 08-MAY-1993
DEFINITION nek1-serine/threonine- and tyrosine-specific protein kinase [mice,
erthroleukemia cells, mRNA, 4263 nt].
ACCESSION  S45828
VERSION    S45828.1  GI:256854
KEYWORDS
SOURCE     Mus sp. erythroleukemia cells.
ORGANISM   Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 4263)
Levin, K., Mizzen, L., Motro, B., Ben-David, Y., Bernstein, A. and
Pawson, T.
TITLE      A mammalian dual specificity protein kinase, Nek1, is related to
cells
the NIMA cell cycle regulator and highly expressed in meiotic germ
JOURNAL    EMBO J. 11 (10), 3521-3531 (1992)
MEDLINE    93010942
REMARK     Genbank staff at the National Library of Medicine created this
entry [NCBI glibseq 114983]. From the original journal article.
This sequence comes from Fig. 1B.
FEATURES
Source
1..4263
/organism="Mus sp."
/db_xref="taxon:10095"
1..4263
/gene="nek1"
/translation="MEKVRVLRKIGEGSGKAVLVKSTEDGRHVRVIREINISMDSKE
ROESREVVALNMKHPNIVQKSEFENGSLVTVMDYCGGDLFRKINAKALFOE
DOILMEFYQICALKHVDKRIIHRDIRSONIFLTKDGTOLDFGIAKAVLNSTVELA
RQICGPVYLSPICENKPYNNKSDIMALGVLELCTIKHAFAGMKMLVLAISG
SPRVSAPHAYDPLSLISOLTERNRDRPSVNSTLEKGFIAKRIEKLSPOLIAEEFC
LTKLSFGQPIPKRRPASGGCVSEFPAOKITTPAKKGVPLTYKKYGDKKLEKPP
PKHKQADLPVKKMSGEERKMSSEPAKRRPLEIEKKKKODORLEIKAKOMKRO
EKKORLEIRAKQGMKRVLRKSGSEYKASFRPTGAVSPSCSPGQCEHHAIED
OMORLEARNKAKGKIGRVLPRQYKALAVENAOVEEFLORREKMAONKARAG
HVVYLRKQIRLQNPNEKQIKAKLSEKAGQIKQZETBELLEKQEDSKQY
TNARAVLKEQLERKREKREKREKWEHLEKAGQIKQZETBELLEKQEDSKQY
KPVISVSALEKVGDLGSLTDQEEEMKSSAISSKRELIRLNEMAKQEDSKRO
HHSQSEVGHKDEREYETENAISSDRKWMGOLVPLDAVTLDTSFSAETKRTVG
EVIRKIDNSPKRWGKNPTDSVLKLGAEALQL."
BASE COUNT 1364 a 829 c 1065 g 1005 t
ORIGIN
alignment_scores:
Quality: 5040.50 Length: 1214
Ratio: 4.496 Gaps: 6
Percent Similarity: 92.339 Percent Identity: 81.796
alignment_block:
US-09-783-320-4 x S45828 ..
Align seg 1/1 to: S45828 from: 1 to: 4263
1 MetGlyLysTyrValArgIleuInLysIleGlyIleuGlySerPheGlyIly 17
|||||
576 ATGGAAGATGATGTGAGACTGCAGAAAGATGTGAGAGAGTTGATTTGAAA 625

```

```

17  sAlaIleLeuValLysSerThrGluAspGlyArgIntYValIleLysG 34
|||||
626 AGCTGCTCTTGTAAATCGACAGAGATGCGACATTAATGTCATCAAGG 675
34  IuIleAsnIleSerArgMetSerLysGluArgGluGlnSerArg 50
|||||
676 AAATTAACATCTCAAGAAATGTCATTAAGAAAGCAAAATCAAGAGGA 725
51  GluValAlaValLeuAlaAsnMetLysHisProAsnIleValGlnTyr 67
|||||
726 GAAGTTCGTGATGCGCAACATGACATCCAAATATGTCATATATA 775
67  GluSerPheGluGluAsnGlySerLeuTyrIleValMetAspTyrCysG 84
776 AGAATCATTTGAGAAATAGGCTCTCTACATAGTAATGATATCTG 825
84  IuGlyGlyAspLeuPheLysArgIleAsnAlaGlnLysGlyValLeuPhe 100
|||||
826 AAGGAGTGATTTGTTAAACGAATAATGCTCAGAAAGCGCTCTGTT 875
101  GlnGluAspGlnIleLeuAspTyrPheValGlnIleCysLeuAlaLeu 117
|||||
876 CAAGAAGACAGATTTGACTGCTGTCAGATATGTTGCTGCTGAA 925
117  sHisValHisAspArgLysIleLeuHisArgAspIleLysSerGlnAsn 134
|||||
926 GCATGTACATGATGAAAAATTTCTCACGAGACATTAAGTCACAGAA 975
134  IspheLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAla 150
976 TATTTCTAACCAACATGCGACATGACGCTTGAGATTTGGAAATGCT 1025
151  ArgValLeuAsnSerThrValGlnLeuAlaArgThrCysIleGlyThrP 167
1026 CAGATTCCTTAATAGTACTGTAGACTGCTGCAATTCGATAGCACTGC 1075
167  oTyrTyrLeuSerProGluIleCysGluAsnLysProTyrAsnAspLys 184
1076 AATACACTGTTGACCTGAAATCTGGAACACCCCTTAATACAAATAAA 1125
184  eAspIleThrAlaLeuGlyCysValLeuTyrGluLeuCysThrLeuLys 200
1126 GTGACATTTGGGCTTTGGCTGTGTCTTAAAGATGTGTATACACTTAA 1175
201  HisAlaPheGluAlaGlySerMetLysAsnLeuValLeuLysIleLeu 217
1176 CATGCATTTGAAGCTGGAACATGAAAAACCTGCTACTGAAATATCTC 1225
217  rGlySerPheProProValSerLeuHisTyrSerTyrAspLeuArgSer 234
1226 CGGATCTTCTCTCCAGTGTCTCCACATTAATCTATGATCTCGCAGCT 1275
234  euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
1276 TGTGTCTCATGATTTTAAAAAATCTCTAGGATAGCATCAGTCAAC 1325
251  SerIleLeuGlnLysGlyPheIleAlaLysArgIleGluLysPheLeu 267
1326 TCCATATTTGAGAAAGTTTATTAACGATTAACGATTAAGTATTTCTC 1375
267  rProGlnLeuIleAlaGlnLysPheCysLeuLysThrPheSerLysPhe 284
1376 CCTCAGCTATTTGCAAGAAATTTGTCTAATAAACACTTTCAAAAGTTG 1425
284  LysGlnProIleProAlaLysArgProAlaSerGlyGlnAsnSerIle 300
1426 GACCACAGCTCTCCAGGTAATAAGACAGCTCAGGCAAGAGGCTGACT 1475
301  SerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyrGly 317
1476 TCTTTTGTCCCTCTCTCAGAAATACAAAGCTGCTGCTAATACGAGGT 1525

```

```

317 eProLeuAlaTyrLysLysTyrGlyAspLysLysLeuHisGluLysLysP 334
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1526 GCCTTTAAATGATTAAGAGTATGAGATTAAGAGTATGAGATTAAGAG 1575
334 rOLeugInLysHisLysGlnAlaHisGlnThrProGluLysArgValAsn 350
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1576 CACCCCAATTAACAGCCCATCAATTCCTGGAAGAAATGAAT 1625
351 ThrGlyGluArgLysLysSerGluAlaAlaArgLysArg 367
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1626 TCTGGAGAGAAAGAAAGAAATGCTGAGGAGCAGCAAAAAAAGAG 1675
367 gLeuGluPheIleGluLysGluLysGluLysAspGlnIleIleSerL 384
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1676 GTTGGATTTTATGAGAAAGAAAGAAAGATCAG..ATTAGGT 1722
384 eUeLysAlaGluGlnMetLysArgGlnLysGluArgLeuGluArg 400
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1723 TCTGGAAGCTGAGCAGATGAAAGCGCAGAGAGAGCGGTTGGAGAG 1772
401 IleAsnArgAlaArgGluGlnGlyTyrPArgAsnValLeuSerAlaGly 417
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1773 ATAAATAGGGCCAGGAAACAAGATGAGAAATGTTTAAAGCGCTGCTG 1822
417 ySerGlyValLysAlaProPheLeuGlySerGlyGlyThrIleAlaP 434
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1823 AAGCGGTAAAGTAAAGCTTCTTTTGGCATTTGGAGGCGCTGCTCTC 1872
434 rOSerSerPheSerSerArgGlyGlnTyrGlnHisIleAlaIlePhe 450
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1873 CATCACCGCTTCTCTCTGAGGCGCAGATGACATTACCATGCCATT 1922
451 AspGlnMetGlnGlnGlnArgAlaGluAspAsnGluAlaLysTyrPArg 467
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1923 GACCAATGCGAGCGGCTAAGAGCAGAAATGAAGCAAGAGAGAGAG 1972
467 gGluIleTyrGlyArgLysLeuProGluArgGlnLysGlnLeuAlaY 484
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1973 GGGATCTATGCTGATGCTCCCAAGAAAGCAAAAGCACTTACGCTG 2022
484 aLIGluArgAlaLysGlnValGluGluPheLeuGlnArgLysArgGluAla 500
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2023 TAGAGAGGCCAACCAAGTGAAATCTCTACAGCTTAAACGAGAGCT 2072
501 MetGlnAsnLysAlaArgAlaGluGlnHisMetValTyrLeuAlaArg 517
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2073 ATGCAAGATTAACCCCGAGCCGAGAGACACGTGTTATTGCGCAAGCT 2122
517 uArgGlnIleArgLeuGlnAsnPheAsnGluArgGlnIleLysAlaL 534
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2123 GAGCAATTAAGACTACAAAATTTTATGAGCGCCCAACAGATTAAAGCCA 2172
534 ySLeuarGlyGluLysLysGluAlaAsnHisSerGluGlnGlnGluY 550
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2173 AACTTCGGTGGAGATTAAGAGCTGATGTGCCAAAGAGCAAGAGCA 2222
551 SerGluGluAlaAspMetArgArgLysLysIleGluSerLeuLysAlaH 567
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2223 ACTGAAGAGACTGACATGAGCGCTCAAAAAGATGAGTCACTTAAAGCGCA 2272
567 sAlaAsnAlaArgAlaValLeuLysGluGlnLeuGluArgLysArgL 584
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2273 AACCAATGCGCTGCTCTCTACTTAAAGAAAGAGCTGAGCGCAAGAA 2322
584 ySGluAlaTyrGluArgGluLysLysValTyrPGLuGlnHisLeuValAla 600
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2323 AGGAAGCTTATGAAAGAGAAAGAAATGATGAGAAAGCAATTTGCGCGCG 2372
601 LysGlyValLysSerSerAspValSerProProLeuGlyGlnHisL 617
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2373 AGG...GTAAAGACTCAGATGTTCTCTGCTTGAAGCTTCTTGAAG 2419
617 rGlyGlySerProSerLysGlnGlnMetArgSerValIleSerValThrS 634
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2420 AGGTGCTTCTCCATCAACAGCAGCAGTGAAGCCTGCTCATTTCTGCACTT 2469
634 eAlaLeuLysGluValGlyValAspSerSerLeuThrAspThrArgGlu 650
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2470 CAGCTTGAAGAAAGTGGCGCTGATGAGAAATTTAACTGATACCAAGAG 2519
651 ThrSerGluGlnTyrThrAsnAlaIleSerSerLysArgGly 667
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2520 .....GAAATGAGAAAGAGTAACTACGTATTTCAAGTAACGAG 2563
667 uIleLeuArgArgLeuAsnGluAsnLeuLysAlaGlnGluAspGlyLysG 684
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2564 AATCTGCTGAGCTTAATGAAATCTTAAGCTCAAGAGATGAAAGG 2613
684 lYMetGlnAsnLeuSerAspThrPheGluIleAsnValHisGluAspAla 700
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2614 AAAAGCAGCATCACTAGCTTCTGTGAGACCGTTGCTCACAAAGATGAG 2663
701 LysGluHisGluLysGluLysSerValSerSerAspArgLysLysTyrPGL 717
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2664 AGAGAGTATGAGACAGAAATGCCATTCTCTGATCCGCAAGAACTGGGA 2713
717 uAlaGlyGlyGlnLeuValIleProLeuAspGluLeuThrLeuAspThrS 734
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2714 GATGGAGGCTCAGCTTGTGATTCCTCTCGATCAGTACAGTGTATACAT 2763
734 eRPheSerThrThrGluArgHisThrValGlyValIleLysLeuGly 750
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2764 CCTTCTCTGCACCGAAAAACATACCTGTGGAGAGCTTATTAATTAAT 2813
751 ProAsnGlySerProArgArgAlaTyrPGLysSerProThrAspSerVa 767
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2814 TCTAATGCTCTCCAGAAAGAGCTGGGGAAGAAACCTCACAGATTCTG 2863
767 lLeuLysIleLeuGlyGluAlaGluLeuGlnLeuGlnThrArgLeuLeuG 784
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2864 GCTGAATATCTTGAGAACCTGATTAACGCTATAGACAGAACACTACAG 2913
784 lAsnThrThrIleArgSerGluIleSerProGluGlyLysLysTyrLys 800
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2914 AAAACATCTTTTAAAGTGAAGTGTATGCTGAAGAGAGAGAAATACAAA 2963
801 ProLeuIleThrGlyLysLysValGlnCysIleSerHisGluIleAs 817
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2964 CCTTACTTACTGAGAGAGAGATTCGACGTCCATTCAAAAAGAAATATAA 3013
817 nProSerAlaIleValAspSerProValGluThrLysSerProGluPheS 834
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3014 TCCATCAGCTACCTGTGATTCT...ACTGAAGCAAGAAAGTCCAAAGTTTA 3060
834 eArgAlaSerProGlnMetSerLeuLysLeuGlnGlyAsnLeuGlnGlu 850
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3061 CTGAGGCTCTCCCAAAATGTCA.....GAGGAAATGTGGAAGAAA 3101
851 ProAspAspLeuGlnThrGluIleLeuGlnGlnGluProSerGlyThrAsn 867
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3102 CCTGATGATTTGGAACAGAGTCTTACAAAGCCAAAGTACAGACACAC 3151
867 sAspGluSerLeuProCysThrIleThrAspValTyrPLeuSerGluL 884
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3152 AGATGGAGTTTCCACTGTTCTTATGATGTGTGAGCTACAGAGAAAGG 3201
884 ySGluThrLysGluThrGlnSerAlaAspArgIleThrIleGlnGluAsn 900
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3202 AAGAGCTTAAGAAAGTGAAGTGAAGAAAGTGTGCTGACAGCAGAT 3251
901 GluValSerGluAspGlyValSerSerThrValAspGlnLeuSerAspI 917
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3252 GAAGTTGTGAAGATGAATTCAGAGGAGCTGAGCAATCTCTGAAGA 3301
917 eHisIleGluProGlyThrAsnAspSerGlnHisSerLysCysAspValA 934
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

3302 TCAGAGAGATCTCGACAGTAGACGATTCCTCCGACTGTGGCTGTGATGTAG 3351
934 splysserValglnprogluprophephistlyValvalhisSerGlu 950
3352 AGAAGTCATGACGACGAGATTCCTCCAGAAAGTGTTCATTCATTCAG 3401
951 HisleuValproglValglnserValglnCysSerProglu 967
3402 GACTTGGACTTA.....GTTCAGGACAGTTCCTCACCAGAGA 3442
967 userPheAlaPheArgSerHisSerHisleuProProlysnlysnLys 984
3443 ACCAATTCGAATCGATCTCCTCATTCCTCCACCAAAACCTAAGACA 3492
984 YssanSerleuValleuValSerThylleuPheAspAlaAsn 1000
3493 AGAATTCCTTACTGATTCGATTCACGCTCTGTTGATGACAACAT 3542
1001 ProlyMetleuArgThylleuSerleuProAspLeuSerlysnPheArg 1017
3543 CCAGAGATGCTGAGACCTGCTCCTCAGATCTTCACAGCTGTTCCAG 3592
1017 gThrleuMetAspValProThylValglnAspValarglnAsnleu 1034
3593 AACCTTATGACGCTTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTG 3642
1034 IuileaspGluileuAspGluAsnIleuGlnGluGlyProSerAspSer 1050
3643 AAATCGATGAGCTGAGATGAACCAATTAAGAGGCTCTCTATTCC 3692
1051 GluAspIleuValPheGluGluThrAspThrAspLeuGlnIleuGlnAl 1067
3693 GAGGACACTGATATTGAAAGAACTGACACAGATTTCACAGACTTCAGG 3742
1067 aSerMetGluIleuValleuArgGluIleuGlnGluGlyIleuSerGlu 1084
3743 CTCATGAGACGCTGCTAGGAGCAACGCTGACGATACGATACGATGAG 3792
1084 IuGlnGluSerValleuLysAsnSerAspValglnProThrlAlaAsnGly 1100
3793 AGGAAGATCTGTTTAAAGACGAGATGTGAGCAGACAGACAGAGAGGG 3842
1101 ThrAspAlaIleuAspGluAspAspAsnProSerSerGluSerAlaLeuAs 1117
3843 ACAGATGCCCGACGAGAGGAGACACCCAGCAGGAGAAAGGCC..CTGAA 3891
1117 nGluGluThrlPheHisSerAspAsnSerAspGluIleuAlaSerGluCysG 1134
3892 CGAGAGATGCGCTCAGATATATAGTGCCTGAGACACAGATGATGTG 3941
1134 IuCysAspSerValPheAsnHisleuGluIleuArgLeuHisleuGlu 1150
3942 AATATGACGCTGCTTAAACATTTAGAGAACTAGACATTCCTGCGAG 3991
1151 GlnGluMetGlyPheGluLysPhePheGluValTyrGluLysIleuLysAl 1167
3992 CAGAAATGGCTTTGAAAGTCTTGTGAGGTTTATAGAAATGAAGG 4041
1167 alenHisGluAspGluAspGluAsnIleuGluIleuCysSerLysIleuValG 1184
4042 TATTCATGAGATGAAGATGAATATGAAATTTGTTCAACAATGTTG 4091
1184 IuAsnIleuGluLysnGluHisGlnHisleuTyrAlaLysIleuHis 1200
4092 AGAATATTTTGGCAATGAGCCAGCATCTCTATGCCAAGATTCGCAT 4141
1201 LeuValMetAlaAspAlaTyrGlnGlnIleuAspAsnAspGlu 1214
4142 TTATGTCATGCGAGATGAGACCTATCAGAGATATATGATGAA 4183

seq_name: gb_pr:AF155113
seq_documentation_block:

```

```

LOCUS AF155113 3555 bp mRNA linear PRI 05-JAN-2000
DEFINITION Homo sapiens NY-REN-55 antigen mRNA, partial cds.
ACCESSION AF155113
VERSION AF155113.1 GI:5360120
KEYWORDS
SOURCE
ORGANISM human.
Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 3555)
Scallan,M.J., Gordan,J.D., Williamson,B., Stockert,E., Bander,N.H.,
Jongeneel,V., Gure,A.O., Jager,D., Jager,E., Knuth,A., Chen,Y.-T.
and Old,L.J.
Antigens recognized by autologous antibody in patients with
renal-cell carcinoma
Int. J. Cancer 83 (4), 456-464 (1999)
JOURNAL
MEDLINE 99438124
PUBMED 10508479
REFERENCE
2 (bases 1 to 3555)
Scallan,M.J., Gordan,J.D., Williamson,B., Stockert,E., Bander,N.H.,
Jongeneel,V., Gure,A.O., Jager,D., Jager,E., Knuth,A., Chen,Y.-T.
and Old,L.J.
Antigens recognized by autologous antibody in patients with
renal-cell carcinoma
Int. J. Cancer 83 (4), 456-464 (1999)
TITLE
Direct Submission
JOURNAL
Submitted (28-MAY-1999) Ludwig Institute, Sloan-Kettering
Institute, 1275 York Ave, New York, NY 10021, USA
FEATURES
source
1..3555
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="1973/10-4"
/cell_type="renal cell carcinoma"
<..2317
/feature="similar to Mus musculus NRK serine/threonine
kinase"
/codon_start=2
/product="NY-REN-55 antigen"
/protein_id="AAD42879.1"
/db_xref="GI:5360121"
/translation="PHYAIFDQMOORAEDEAKKREIYGRGLPERQGGOLAVERA
KQEPRIORERAKMOKRABGHVYLLARLQRLQFNRROOIKAKLGEKKENHS
EKGESERADMRKREIKANANAARAVLEQERKRAYREKKWEHILVANKV
KSDVSPFPGCHQETGSPGKONSYSVTSALKEVDSLSLDTRETSSEMOKTNNA
ISKRILRLRLNENKQEDKQKONSDPPEINVEDAKHEKESVSDDKMEAG
GOLVPLDLTLIDPSTPEPRKYPITGCEKYOCISHEINPSAIYDPSVETKSGEALO
LOVELENTITSEISPRGEKYPITGCEKYOCISHEINPSAIYDPSVETKSGEALO
LITOVQSVQCSPESEAFRSHSLPFRKNNKSLKLGSLGPIRANNDKMLFQCSL
PDLKLFRLIMVPTVGVRODNLIDLEEDNKISPSDSDEDDNRSSEALNEHNSNS
SMOILREOPGSEYSSEESVLKNSDVEPTAGNDVDEDDNRSSEALNEHNSNS
DGEIASECCDVSFNLIELRLHLEQEGEFKEFEYEKTAIHEDENITELSKIV
ONTLENEHOLYAKLILHVMADGAYQEDNDE"
BASE COUNT 1247 a 572 c 728 g 1008 t
ORIGIN
alignment_scores:
Quality: 3945.00 Length: 771
Ratio: 5.130 Gaps: 0
Percent Similarity: 99.741 Percent Identity: 99.611
alignment_block:
US-09-783-320-4 x AF155113
Align seg 1/1 to: AF155113 from: 1 to: 3555
444 GlnHisTyrHisAlaIlePheAspGlnMetGlnGlnGlnTyrGlnIleuAs 460
|||||
2 GAACATTTACATGCCATTTTTCACCAATGCGACGACCAAGAGAGAGA 51
460 PaangluAlaLysTTPPLysArgGluIleTyrGlyArgGlyIleuProglu 477
|||||
52 TAATGAAGCTAAATGAGAAAGAAATATATGCTGAGGCTCTTCCAGAA 101

```

477 rglInlysglInleuAlaValgluArGAlaLysglInValgluIuphe 493
 102 GGCAAAAGGCGACAGCTAGCTGTACAAAGAGCTTAACAGTAGAAGAGTTTC 151
 494 LeuGlnArGlyArGluAlaMetGlnAsnLysAlaArGAlaGluIylH 510
 152 CTGCGAGCGAAAAGGAAAGCTATGACAAATTAAGCTGAGCCGACAGACA 201
 510 sMeValIyrlEuAlaArGleuArGlnIleArGleuGlnAspPheAsnG 527
 202 TATGTTTATCTGCGAAGACTGAGGCAAAATTAACACTACAAATTTCAAG 251
 527 luArGlnGlnIleLysAlaLysLeuArGlygluLysgluAlaAsn 543
 252 AGCGCAACAGATTAAAGCCAAACTCTGTGTGAAAGAAAGAAAGCTTAAT 301
 544 HisSerGluGlnGlnGlySerGluGluAlaAspMetArGArGlyLys 560
 302 CATCTGAAAGAGCAAGAAAGAGAGCTGACACTGAGCGCAAAAA 351
 560 sIlleGluSerLeuLysAlaHisAlaAsnAlaArGAlaAlaValleuLysG 577
 352 AATCGAATCACTGAAGGCCCATGCAAAATGCAGCTGCTGTACTTAAAG 401
 577 luGlnLeuGluArGlyArGlysgluAlaTyrgluArGluLysLysVal 593
 402 AACAACTGAACGAAAGAAAGAGAGCTTATGAGAGAGAAAAAAGTG 451
 594 TrpGluGluHisLeuValAlaLysGlyValLysSerSerAspValSerP 610
 452 TGGGAAGAGCATTTGGTGGCTTAAAGAGTTAAAGTTCTGATTTCTCC 501
 610 oProleuGlyGlnHisGluTrpGlyLysSerProSerLysGlnIleMetA 627
 502 ACCTTTGGACAGCATGAACAGGTGGCTCTCATCAAAAGCAACAAATGA 551
 627 rGSerValIleSerValThrSerAlaLeuLysGluValGlyValAspSer 643
 552 GATCTGATTTCTGACTCTGACTTGAAGAGAGTTGGCTGGAGACAG 601
 644 SerLeuThrAspThrArGluThrSerGluGluMetGlnLysThrAsnAs 660
 602 AGTTTAACCTGATACCCGGGAAACTTCAGAAAGAGATGCAAAAAGCAACA 651
 660 nAlaIleSerSerLysArGluIleLeuArGArGleuAsnGluAsnLeu 677
 652 TGCTATTTCAGTAAGCGAGAAATACCTCGTAGATTAAATGAATAACTTA 701
 677 ySAlaGlnGluAspGluLysGlyMetGlnAsnLeuSerAspThrPheGlu 693
 702 AAGCTCAGAGATGAAAAAGAAAGCAAGATCTCTGTACTTTTGGAG 751
 694 IleAsnValHisGluAspAlaLysGluHisGlnLysGluLysSerValSe 710
 752 ATAATGTTCATGAAGATGCCAAAGAGCATGAAAAAGAAAAAACACTTTC 801
 710 rSerAspArGlyLysTrpGluAlaGlygluLysGlnLeuValIleProLeuA 727
 802 ATCTGATCGCAAGAGTGGAGGCGAGGCTGATGATTCCTCTCTGG 851
 727 spGluLeuThrLeuAspThrSerPheSerThrThrGluArGHisThrVal 743
 852 ATGAGTTAACACTAGATACATCTTCTCTCACTGAAAGCACTAACAGAG 901
 744 GlyGluValIleLysLeuGlyProAsnGlySerProArGArGAlaTrpGl 760
 902 GGAGAAAGTTATTAAATAGTCTTAATGATCTCCAAAGAGAGACCTGGGG 951
 760 yLysSerProThrAspSerValLeuLysIleLeuGlnGluAlaGluLeuG 777
 952 GAAAAGTCCGACAGATTCTGTCTTAAAGATACCTTGGAGAGCTGAACCTAC 1001

777 luLeuGlnThrGluLeuLeuGluAsnThrThrIleArGSerGluIleSer 793
 1002 AACTTCAGACAGAACTATTAGAAAAATACAACTATTAGAAAGTGACATTCT 1051
 794 ProGluGlygluLysTrpLysProLeuIleThrGlygluLysValG 810
 1052 CCGAAGGGGAAAAAGTACAAACCTTAATTACTGGAGAAAAAAGATGACA 1101
 810 nCysIleSerHisGluIleAsnProSerAlaIleValAspSerProValG 827
 1102 ATGATTTTCATGAATTAACCCATACAGCTTGTGTTGATCTCTCTGTTG 1151
 827 luThrLysSerProGluPheSerGluAlaSerProGlnMetSerLeuLys 843
 1152 AGACAAAAGTCCCGAGTTCAAGAGCAATCTCCACAGATGATCATTTGANA 1201
 844 LeuGluGlyAsnLeuGluGluProAspAspLeuGluThrGluIleLeuG 860
 1202 CTGGAAGAAATTTAGAAAGACCTGATGATTTGAAAGAAAGAAATTTACA 1251
 860 nGluProSerGlyThrAsnLysAspGluSerLeuProCysThrIleThra 877
 1252 AGAGCCAAAGTGAACAAAGAGATGAGCTTGCATGCACTATTACTG 1301
 877 spValTrpIleSerGluGluLysGluThrLysGluThrGlnSerAlaAsp 893
 1302 ATGTGTGATTACTGAGGAAAAAGAAACAAAGAAACACAGTCGCGCAGAT 1351
 894 ArgIleThrIleGlnGluAsnGluValSerGluAspGlyValSerSerTh 910
 1352 AGGATCACCATTCAGAAATAATGAAGTTCTGAAGATGAGTCTCGAGTAC 1401
 910 rValAspGlnLeuSerAspIleHisIleGluProGlyThrAsnAspSerG 927
 1402 TGTGGACCAACTTATGACATTCATATAGAGCTTGGAACCAATGATTTCT 1451
 927 lnhSerLysCysAspValAspLysSerValGlnProGluProPhePhe 943
 1452 AGCACTCTTAATGTGATGTAGATAGCTGTGTCAACCGGAACCATTTTTC 1501
 944 HisLysValValHisSerGluHisLeuAsnLeuValProGlnValGlnSe 960
 1502 CATAAAGTGGTTCAATTCAGAACACTTGAACCTTAGTCCCTCAAGTTCAATTC 1551
 960 rValGlnCysSerProGluGluSerPheAlaPheArGSerHisSerHisL 977
 1552 AGTTCAAGTGTTCACCAAGAAATCCCTTTCATTTGATCTCACTCGCATTT 1601
 977 euProProLysAsnLysAsnLysAsnSerLeuLeuIleGlyLeuSerThr 993
 1602 TACCACCAAAAAATTAATAACAGAAATTCCTGCTGATTTGAGACTTTCACCT 1651
 994 GlysLeuPheAspAlaAsnAsnProLysMetLeuArGThrCysSerLeuPr 1010
 1652 GGTCTGTTTATGCAACAACCAACCAAGATGTTAAGACATGTCTCACTTCC 1701
 1010 oAspLeuSerLysLeuPheArGThrLeuMetAspValProThrValGlyA 1027
 1702 AGATCTCTCAAGCTGTGTGAAACCTTATGATGTCTCCACAGCTGAGAG 1751
 1027 spValArGlnAspAsnLeuGluIleAspGluIleLysAspGluAsnIle 1043
 1752 ATGTTCTCTCAAGACAATCTTGAATAATGATGAATAATGAAGATCAAAACATTT 1801
 1044 LysGluGlyProSerAspSerGluAspIleValPheGlnGluThrAspTh 1060
 1802 AAAAGAAAGACCTTCTGATCTGAAAGCATTTGTTGAAAGAAACGAGAC 1851
 1060 rAspLeuGlnGluLeuGlnAlaSerMetGluLysLeuLeuArGlyGlnInp 1077
 1852 AGATTTCAGAGACTCGAGGCTCGATGTGAACAGATTACTTAGGGAACAC 1901
 1077 roGlyGluGluLysrSerGluGluGluSerValLeuLysAsnSerAsp 1093


```

|||||
503 GCTGATGTGACCTTCACACTGCTGTGTGATGCAACACCAACCAACATGCT 552
1004 eUArTThCysSerLeuProAspLeuSerLysLeuPheArqThrLeuMet 1020
553 TTAGGACATGTCTACCTCCAGATCTCTCAAGCGTTCACAGACCCCTTAG 602
1021 AsPValProThrValIGLAspValArgGLAspAsnLeuLysIleAspGI 1037
603 GATGTTCACCGTAGAGATGCTTCAGACAGACATCTTGAATATAGATGA 652
1037 uLleLysAspGIuAsnIleLysGIuLysProSerAspSerGIuAspIle 1054
653 AATTGAAGATGAAACATTTAAGAGACCTTCTGATTCGAGACATG 702
1054 aPhGGLuGIuThrAspThraspleuGIuLeuGIuLeuAlaSerMetGIu 1070
703 TGTTTGAAGAACTGACACAGATTTACAGAGCTGCGAGGCTCGATGGA 752
1071 GluLeuLeuArgGIuGIuInProGIuGIuLysSerGIuGIuLysGIu 1087
753 CAGTACTTAGGGAACAACCTGCTGAAGAAATACAGTGAGAGAGAGATG 802
1087 rValLeuLysAsnSerAspValIGLupProThrAlaAsnGIyThrasPVal 1104
803 AGCTTTGAAGAAACAGTGTGTGGAGCCCACTGCAAAATGGGACAGATG 852
1104 laAspGIuAspAspAsnProSerSerGIuSerAlaLeuAsnGIuGIuTrp 1120
853 CAGATGAGATGACAAATCCAGCAGTGAAAGTGCCCTGAGAGAAATG 902
1121 HisSerAspAsnSerAspGIuGIuIleAlaSerGIuGIuGIuGIuAsp 1137
903 CACTCAGATTAACAGTGAATGCTGAATTCGTAGTGAATGCTAATGCCAT 952
1137 rValPheAsnHisLeuGIuGIuLeuArgLeuHisLeuGIuGIuGIuMet 1154
953 TGTCTTTAACCTTTAGAGAACTGAGACCTTCATCTGAGACAGAAATG 1002
1154 LysPheGIuLysPhePheGIuValTyrGIuLysIleLysAlaIleHisGIu 1170
1003 GCTTTGAAAAATCTTTGAGGTTATGAGAAATAAGGCTATCTATGAA 1052
1171 AspGIuAspGIuAsnIleGIuLysSerLysIleValGIuAsnIleLe 1187
1053 GATGAAGATGAAATATGAAATTTGTTCAAAATAGTCAAAATATATTT 1102
1187 uGIyAsnGIuHisLeuTyrAlaLysIleLeuHisIleLeuValMet 1204
1103 GGGAAATGACATCGCATCTTTATGCCAAGATCTTCATTAGTCATCG 1152
1204 laAspGIuAlaTyrGIuGIuAspAsnAspGIu 1214
1153 CAGATGGAGCCTACCAAGACATATGATGAA 1184

```

seq_name: gb_pr:AK027580

seq_documentation_block:

LOCUS AK027580 2975 bp mRNA linear PRI 15-MAY-2001

DEFINITION Homo sapiens CDNA FLJ14674 fls, clone NT2RP2003912, highly similar

TO SERINE/THREONINE-PROTEIN KINASE NEK1 (EC 2.7.1.-).

ACCESSION AK027580

VERSION AK027580.1 GI:14042352

KEYWORDS oligo capping; fls (full insert sequence).

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (bases)

AUTHORS Isegai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ishii,S., Kawai,Y., Salto,K., Yamamoto,J.,

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
Location/Qualifiers
1..2975
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RP2003912"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_idb="NT2RP2"
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction."
354..1316
/codon="unnamed protein product"
/protein_id="BAB5209.1"
/db_xref="GI:14042353"
/translation="MKLVLRITSGSFPPVSLHYSDLSLVSOLFKNRPDRPSVNS
ILEGFIAPKREKFLSPQIAEEPCFKTSKSGSDIPAKRRASGONSISVPAOKIT
KPAKVGIPLAYKYGDKKLHEKKPILOKHQAHQPEKRVNTEGERRKISEEAAKRR
LEPEKKKKODITSLMKAQMKROEKRELRERINRAREGVNLSAGSGGEVYKAP
LGSGGTAPSPSFRGQYEHVHAPMOQOQRAEDKAKREIYIGRLPERGLIPGV
RPGPYRAGHHHPDMDIRTKIKLTKAVSKOANNRIDCDP"
BASE COUNT
894 a 556 c 562 g 963 t
ORIGIN
alignment_scores:
Quality: 2012.00 Length: 387
Ratio: 5.199 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-783-320-4 x AK027580 ..
Align seg 1/1 to: AK027580 from: 1 to: 2975
91 ArgIleAsnAlaGlnLysGIuValLeuPheGlnGIuAspGIuIleLys 107
3 CGAATAATGCTCAGAAAGCGTTTGTTCAGAGATCGATGATTTGGA 52
107 pTppPheValGlnIleCysLeuAlaLeuLysHisValHisAspArgLys 124
53 CTGGTTTGACAGATATGTTTGCCCTGAAACATGATGATGAAAAA 102
124 leuLeuHisArgAspIleLysSerGIuAsnIlePheLeuThryAspGI 140
103 TTCTTCATCGACACATTAATCTCAGAACATVTTTTTAACTAAGATGA 152
141 ThrValGlnLeuGIuAspPheGIuIleAlaArgValLeuAsnSerThr 157
153 ACAGTACAACTTGAGATTTTGGAAATGCTAGAGTTCTTAATAGACTGT 202
157 lGIuLeuAlaArgThryCysIleGIyThryProTyrTyrLeuSerProGI 174
203 AGAGCTGCGTCGAACTTGACATAGAGGCCCACTACTAGTTCACCTGAA 252
174 leCysGIuAsnLysProTyrAsnAsnLysSerAspIleTrpAlaLeuGI 190

```

```

|||||
253 TTCGTAAGAAACCTTACATATATAAAGTGACATTGGCTGGGG 302
191 CysValLeuTyrGluLeuCysThrLeuYsHisAlaPheGluAlaGlyse 207
303 TGTCCTTTATGACCTGTGTAACCTTAAACATCTTTGACGTGGCAG 352
207 rMetLysAsnLeuValLeuYsHisLeuSerGlySerPheProProValS 224
353 TATGTAACAACTGGTACTGATGATATATCTGATCTTTCCACACCTGT 402
224 erLeuHisTyrSerTyrAspLeuAqSerLeuValSerGlnLeuPheLys 240
403 CTTCGATATATCTCTATGATCTCCACAGTTGGTGTGCTCACTTTTAA 452
241 ArgAsnProArgAspArgProSerValAsnSerLileuGluLysGlyph 257
453 AGAAATCTTAAGGATGACCATCATCACTCCATATTGGAGAAAGGTTT 502
257 eLleAlaLysArgLileGluLysPheLeuSerProGlnLeuAlaGluG 274
503 TATAGCCAAACGATGAAAGTTCTCTCTCTCAGCTTATTCACAGAG 552
274 LuPheCysLeuLysThrPheSerLysPheGlySerGlnProLleProAla 290
553 AATTTGTCTAAACATTTTCGAAAGTTTGATCATACACCTCATATCCAGCT 602
291 LysArgProAlaSerGlyGlnAsnSerLileSerValMetProAlaGlnL 307
603 AAAAGCCAGCTTCAGAGCAAACTCGATTTCTGTATTCGCTGCAAGAA 652
307 sLleThrLysProAlaLalaLysTyrGlyLleProLeuAlaTyrLysLST 324
653 AATTACAAAGCTCCGCTTAATATGAAATACCTTACATATTAAGAAAT 702
324 YrGlyAspLysLysLeuHisGluLysLysProLeuGlnLysHisLysGln 340
703 ATGGAGATATAAAATTTACAGAAAGAAACCACTGCCAAACATATAAGAG 752
341 AlaHisGlnThrProGluLysArgValAsnThrGlyGluLysArgGly 357
753 GCCCATCAACTCCAGAGAAAGAGTGAATACCTGGAGAAAGAGGAGAA 802
357 sLleSerGluGluAlaLalaLysArgArgLeuGluPheLleGluLysG 374
803 AATATCTGAGAGACGACAGAAAGAGAGAGCTGGAAATTTATTGAAAAG 852
374 LuLysLysGlnLysAspGlnLleLleSerLeuMetLysAlaGluGlnMet 390
853 AAAAGAAACAAAGAGATCAGATTATGATTATGAAAGGCTGAAACAAATG 902
391 LysArgGlnGluLysGluArgLeuGluArgLleAsnArgLlaArgGluG 407
903 AAAAGCAAGAAAGAAAGAGCTGGAAAGATAATATAGGGCCAGGAGAA 952
407 nGlyTrrArgAsnValLeuSerAlaGlyLysGlyGluValLysAlaP 424
953 AGGATGAGAAATGTGTAAGTGTGGTGGAGAGTGAAGTGAAGTGAAGCTC 1002
424 rPheLeuGlySerGlyGlyThrLleAlaProSerSerPheSerSerArg 440
1003 CTTTCTGGGCACTGAGAGGACTATAGCTCATCATCTTTTCTTCGGA 1052
441 GlyGlnTyrGluHisTyrHisAlaLlePheAspGlnMetGlnGlnInar 457
1053 GGCAGATATGACATTACCATCTTTTGGACCAATGCGAGCAACAAAG 1102
457 gAlaGluAspAsnGluAlaLysTrrLysArgGluLleTyrGlyLysGlyL 474
1103 AGCGAAGATTAATGAAGCTAATATGAAAGAAAGAAATATATATGTCGAGGTC 1152
474 eupProGluArg 477
|||||

```

```

1153 TTCGTAAGAAAGA 1163
seq_name: gb_pat:AX166554
seq_documentation_block:
LOCUS AX166554 1947 bp DNA Linear PAT 22-JUN-2001
DEFINITION Sequence 45 from Patent WO0138503.
ACCESSION AX166554
VERSION AX166554.1 GI:14546899
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1947)
Plozman,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R.,
Flanagan,P. and Clary,D.S.
Novel human protein kinases and protein kinase-like enzymes
Patent: WO 0138503-A 45 31-MAY-2001;
JOURNAL
Sugen, Inc. (US)
FEATURES
Location/Qualifiers
source 1..1947
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 638 a 357 c 455 g 497 t
ORIGIN
alignment_scores:
Quality: 1025.00 Length: 818
Ratio: 2.293 Gaps: 25
Percent Similarity: 54.645 Percent Identity: 32.763
alignment_block:
US-09-783-320-4 x AX166554
Align seg 1/1 to: AX166554 from: 1 to: 1947
1 MetGluLysTyrValArgLeuGlnLysLileGlyGluGlySerPheGly 17
|||||
1 ATGGATTAAGTACGATGATTAAGCCATGCGGCAAGTGCCCTTGGGAA 50
|||||
17 sAlaLleLeuValLysSerThrGluAspGlyAagGlnTyrValLileLysG 34
|||||
51 ACCATACCTTAGCTAAAGGAAATCAGATACAGAAACAGAGCTTCAAGAAA 100
|||||
34 LuLleAsnLleSerArgMetSerSerLysGluArgGluGlnLysArgArg 50
|||||
101 AGATCAATTTTGAAGAAAGATGCCCATACAGAAAGAAAGAGCTTCAAGAAA 150
|||||
51 GluValAlaValLeuAlaAsnMetLysHisProAsnLleValGlnTyrAr 67
|||||
151 GAAGTGATTCCTCTGAAAGATGAAACATCCCAACATTTGTACCTTCTT 200
|||||
67 gGluSerPheGluGluAsnGlySerLeuTyrLleValMetAspTyrCysG 84
|||||
201 CATTCATTTCAAGAAATGCGAGCGTGTATATGTAATGAAATATTTGG 250
|||||
84 LuLysLysAspLeuPheLysArgLleAsnAlaGlnLysGlyValLlePhe 100
|||||
251 ATGAGAGGAGATCTCATGAAAGAGATCATAGACAACGCGGTGTATATT 300
|||||
101 GlnGluAspGlnLleLeuAspTrrPheValGlnLleCysLeuAlaLeuY 117
|||||
301 AGTGAATTCAGATCTCCGCTGTGTGTACAGATTTCTCTAGAGCTAAA 350
|||||
117 sHisValHisAspArgLysLileuHisArgAspLileLysSerGlnAsn 134
|||||
351 ACATATTCAGACAGAGAAATTTACACAGGACATAAAGCTCAGAGACA 400
|||||
134 lPheLeuThrLysAspGlyThrVal...GlnLeuGlyAspPheGlyLle 149
|||||
401 TTTTCTTACAAAGAGATGTGGCAAGCTTGGGACTTGGTATA 450
|||||

```

150 AlaArgValLeuAsnSerThrValGluLeuAlaArgThrCysIleGlyTh 166
 451 GCAAGAGCTGTAATATTCATGCACTGCGAACTGTATTCGAAAC 500
 166 rProTyrLeuSerProGluIleCysGluAsnLysProTyrAsnAspL 183
 501 ACCCTACTACCTGTCCCGAGAGATCTGTCAAGATAAACCTTACACATA 550
 183 ySerAspIleTyrAlaLeuGlyCysValLeuYrGluLeuCysThrLeu 199
 551 AAACGATATTTGGCTCTGTGGCTGTCTATATGAGCTCTCACAACCT 600
 200 LysHISAlaPheGluAlaGlySerMetLysAsnLeuValLeuLysIle 216
 601 AAACATCTTTGAGGTAACACCTTACAGCAGCTGTCTGTGAAGATTGG 650
 216 eSerGlySerPheProProValSerLeuHisTyrSerTyrAspLeuArg 233
 651 TCACAGCATTCTTCCCAATATCTCGGGGTTTCTCGAGCTCCATT 700
 233 eTLeuValSerGlnLeuPheLysArgAsnProArgAspArgProSerVal 249
 701 CCTGTATCTCAGCTCTTTCAGATATCTCGAGACCGACCATCCATA 750
 250 AsnSerIleLeuGluLysGlyPheIleAlaLysArgIleGluLysPhe 266
 751 AATTCATTTTGAAGAAGCCCTTTTGAAGATCTTATCCCAATATTT 800
 266 uSerProGln.....LeuIleAlaGluLup 275
 801 GACTCTCAGGTAAGTTTGAGTGAGCTGTTGGATTGTCGACAGATT 850
 275 heCysLeuLysThrPheSerLys.....PheGlySerGlnProIle 288
 851 TGGTGCGAGGCTCTGACACGCTGTCTGTAGGTGAGCCACAG 900
 289 ..ProAla.....LysArgProAlaSerGlyLysAsn 299
 901 CACCCAGCCTGATATGTTTAAACATCCCTGTTTCTTCTGTA 950
 299 rIleSerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyrG 316
 951 GATTAATAATGATAGAAAGACCCAAATTCGTCTGTGACATTATG 1000
 316 yIleProLeuAlaTyrLysLysTyrGlyAspLysLysLeuHisGluLys 332
 1001 ATTATTATTATGCTCACTGATGCTGAGAGAGAGACCCAC 1044
 333 LysProLeuGlnLysHisLysGlnAlaHisGlnThrProGluLysArgVal 349
 1045 AAACCA.....AGTTATCACCCATTCTCTCAAGAA 1074
 349 LAsnThrGlyLysGluArgArgLysIleSerGluGluAlaAlaArgLys 366
 1075 .AAATCTGAGTGG 1089
 366 rGArgLeuGluPheIleGluLysGluLysLysGlnLysAspGlnIleIle 382
 1090GATTACGCTCAGAAACGAGCATGGTCCATCCCA 1125
 383 SerLeuMetLysAlaGluGlnMetLysArgGlnLysGluArgLeuGln 399
 1126 AGTCAATAGCCTGCTGATACCTTCAGAGAAATTTGAAGCTCAACATA 1175
 399 uArgLysAsnArgAlaArgGlnGlnGlyTyrPargAsnValLeuSerAlaG 416
 1176 TAAAGTTGAAGTGAGAGCAATGGGCTTCG 1209
 416 LysGlySerGlyGluValLysAlaProPheLeuGlySerGlyGlyThrIle 432
 1209 1209

433 AlaProSerSerPheSerSerArgGlyGlnTyrGlnLysHisAlaIle 449
 1210CCATCTTCTCCGAGCCCAAT 1230
 449 ePheAspGlnMetGlnGlnArgAlaGluAspAsnGluAlaLysTyrL 466
 1231 .TACAAACAGACAGACCTTAAGAAAGTAATGAGAGAGAGCCCTGATTC 1279
 466 ySarGlyGluIleTyrGlyArgGlyLeuProGluArgGlnLysGlyLeu 482
 1280 AGCAG.....CTCCATTATGAAAAAGAAATGAG 1311
 483 AlaValGluArgAlaLysGlnValGluGluPheLeuGlnArgLysArgG 499
 1312 GAACAGAAATTTGGAAGCAGTTAGAGAA 1341
 499 uAlaMetGlnAsnLysAlaArgAlaGluGlyHisMetValTyrLeuAla 516
 1341 1341
 516 rGLeuArgGlnIleArgLeuGlnAsnPheAsnGluArgGlnIleLys 532
 1342ATACCCCAACAGTACCAACATGACATGAAGAAATTAAGA 1380
 533 AlaLysLeuArgGlyGluLysGluAlaAsnHisSerGluGlyGlnI 549
 1381 AAGAAATGGGAGAGACAGACAGACATT 1410
 549 uGlySerGluGluAlaAspMetArgArgLysLysIleGluSerLeuLys 566
 1411GAAAGACTTGAAACAAATGAGCGTTGAGAAACAAAG 1449
 566 LAsnAlaAsnAlaArgAlaAlaValLeuLysGluGlnLeuGluArgLys 582
 1450GAAAGTAA 1458
 583 ArgLysGluAlaTyrGluArgLysLysValTyrGluGlnHisLeuVal 599
 1459 AATCCAA.....CAGAAATTAATAAGC 1481
 599 LAlaLysGlyValLysSerSerAspValSerProProLeuGlnHisG 616
 1482 TAAGAAAGGGGTAAATTT 1500
 616 LuthrGlySerProSerLysGlnGlnMetArgSerValIleSerVal 632
 1500 1500
 633 ThrSerAlaLeuLysGluValGlyValAspSerSerLeuThrAspThrAr 649
 1501GAATTAATTAGCAAAATGTATTTCGAT 1530
 649 gGluThrSerGluGluMetGlnLysThrAsnAlaIleSerSerLys 666
 1530 1530
 666 rGgluIleLeuArgArgLeuAsnGluAsnLeuLysAlaGlnGluAspGlu 682
 1531GAAACATCTCCAGAGAGAGAGCA 1557
 683 LysGlyMetGlnAsnLeuSerAspThrPheGlu.....IleAs 695
 1558 ATGATATTAACCAATATGAACCTTGGAGTGGATGCAATGAATTA 1607
 695 nValHisGluAspAlaLysGluHis.....G 704
 1608 GGAATATGAATGTATAAGAGCATGAGATTAACAGACAAACGATTGG 1657
 704 LysLys.....GluLysSerValSer..... 711
 1658 AAAAATTCACCTGCCAGAAAGCAGGTTTCCACGCAACATCTACCTCT 1707
 712AspArgLysLysTyrGluAlaGlyGly.....GlnLeuValIlePr 725


```

1708 GTGGGAAAGAGGAGCAGTGGATGAGAGCAGCCCTCAGACCTGCGTCA 1757
      :::::::::::::::::::::
725 oluAaPgluLeuThrLeuAspThrSerPheSerThrThrGluArgHisT 742
      :::::::::::::::::::::
1758 GATGATGGGAGGCGGACATCATCCTCCAGCTGCCCGGCGGCTGACA 1807
      :::::::::::::::::::::
742 hrvAlaGlyGluValIleLeuLeuGlyProAsnGlySerProArgArgAla 758
      :::::::::::::::::::::
1808 AT...GGCCAAGTATTGTGATTGAAGCATCCAGAAACAGAAACAG 1854
      :::::::::::::::::::::
759 TrpGlyLysSerProThrAspSerValLeuLysIleLeuGlyGluAlaG 775
      :::::::::::::::::::::
1855 TGGCGGCAATGAGCTCCAGAACTTAAATGATGTCTTTGGCAGACAGACA 1904
      :::::::::::::::::::::
775 ulou 776
      :::::::::::::::::::::
1905 TCTA 1908

seq_name: gb_pat:AX339102
seq_documentation_block:
LOCUS AX339102 654 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 65 from Patent WO0192525.
ACCESSION AX339102
VERSION AX339102.1 GI:18129181
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (sites) Harlocker, S.L., Wang, T., Bangur, C.S., Klee, J.I. and Switzer, A.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0192525-A 65 06-DEC-2001;
CORIXA CORPORATION (US)
FEATURES
source location/Qualifiers
1..654
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 171 a 134 c 121 g 228 t
ORIGIN
alignment_scores:
Quality: 1016.00 Length: 215
Ratio: 4.861 Gaps: 0
Percent Similarity: 97.209 Percent Identity: 95.349
alignment_block:
US-09-783-320-4 x AX339102/rev ..
Align seg 1/1 to reverse of: AX339102 from: 1 to: 654
157 valGluLeuAlaArgThrCysIleGlyThrProTyrtYrLeuSerProG 173
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
643 CTGGAAGAGGGGGCGAGCTGATGAGGGCCCAATCTACTGTCACCTGA 594
173 uileCysGluAsnLysProTyrtYrAsnAsnLysSerAspIleTrpAlaLeu 190
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
593 AATCTGTGAACAAACT.TACAAATAATAAAG.GACATTGGGCTCTGG 546
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
190 lYcYValLeuTyrgluLeuCysThrLeuLysHisAlaPheGluAlaGly 206
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
545 GGGTCTCTTATGAGCTGTACACTTAAACATGCTTTGAAGCTGGC 496
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
207 SerMetLysAsnLeuValLeuLysIleLeuSerGlySerPheProProVa 223
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
495 AGATGAAAACCTGTACTGAAGTAAATATCTGGATTTTCCACCTGT 446
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
223 lserLeuHisTyrtYrAspLeuArgSerLeuValSerGluLeuPheL 240
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
445 GTCTTGCAATTATTCCTATGATCTCCGAGTTTGGTGTCAGTTATTTA 396
240 ysArGAsnProAlaArgAspArgProSerValAsnSerIleLeuGluLysGly 256
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
395 AAGAAATCTTAGGATTAACCATCATGCACTCATATTGGAAAGGT 346
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
257 PheIleAlaLysArgIleGlyLysPheLeuSerProGluLeuIleAlaG 273
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
345 TTTATACCAACACCATTAAGAAAGTTTCTCTCTCAGCTTATTGACAGA 296
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
273 uGluPheCysLeuLysThrPheSerLysPheGlySerGlnProIleProa 290
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
295 AGAATTTGTCTAAAAACATTTTCGAAGTTTGATGACACAGCCATACAG 246
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
290 lAlYsArgProAlaSerGlyGlnAsnSerIleSerValMetProAlaGln 306
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
245 CTAAGAAGACAGCTTCAGCAAAACACTCATTTGTTATGCTGCTCAG 196
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
307 LysIleThrLysProAlaAlaLysTyrglyIleProLeuAlaTyrtYrLys 323
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
195 AAAATTCACAAAGCCTGCCCTAAATATGAAATACCTTTAGCATATTAAGA 146
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
323 sTyrtYrAspLysLysLeuHisGluLysLysProLeuGlnLysHisLysG 340
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
145 ATATGAGAGATTAATAATTACACGAAAGAACACACTGCAGAAACATTAAC 96
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
340 lAlaHisGlnThrProGluLysArgValAsnThrGlyGluGluArgArg 356
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
95 AGCCCATCAAACTCCAGAGAGAGAGTGAATCTGGAGAAAGAGAGG 46
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
357 LysIleSerGluGluAlaAlaArgLysArgArgLeuGluPheIle 371
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
45 AAAATATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

seq_name: gb_dr:BC019916
seq_documentation_block:
LOCUS BC019916 2134 bp mRNA linear PRI 22-JAN-2002
DEFINITION Homo sapiens, clone MGC:29949 IMAGE:5000918, mRNA, complete cds.
ACCESSION BC019916
VERSION BC019916.1 GI:18043742
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 2134) Strausberg, R.
JOURNAL Direct Submission
Submitted (19-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contract: amgdcn.tmc.edu
Guanatane, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M.,
Yoon, V.S., Kovis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ULNL at: http://image.llnl.gov
Series: IRAX Plate: 42 Row: 1 Column: 10
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not

```

Identify to protein.

FEATURES
source

Location/Qualifiers
1..2134
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:29949 IMAGE:5000918"
/tissue_type="Placenta, choriocarcinoma"
/clone_lib="NH.MGC_10"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"
272..1624
/product="Unknown (protein for MGC:29949)"
/protein_id="AAH19916.1"
/db_xref="GI:18043743"
/translation="MKHPHIVAFKESFEAEHLYIMETCDGDLMOKITOOKKLEP
EDKILMFYOMKLGYNHJHKKVHLHDKSNIFLJONGKVLGDFGSAKLLSPNAF
ACTYVGTPTYPEPEIWMENLYNNKSDIYELCTLKHPRQANWNLKVCQ
GCISPLSHSYELOLVKOMKRNPSATSLSGIVARLVORCLPEIIMEVG
EVLLEIKSKHNTPRKTPSRIRIALGNBSAVQEEEDPKGSHDLSEINELVE
SALRRVNRKGNKSNVHLKASSPNLHRRQKVNPTALALNASILTSLSLAEDD
RGGSVLYKSKNTKRWMLKETPDILNLKADISLAFQYTTIYRPGSEGLKPLSE
ETEASDSVDGHDVYILDPERLEPLDEEDTDPEEDDNPDMVSELKRRAGWGLCDR
"

CDS

BASE COUNT 678 a 416 c 494 g 546 t
ORIGIN

Alignment_scores:
Quality: 882.00 Length: 412
Ratio: 3.031 Gaps: 6
Percent Similarity: 70.631 Percent Identity: 42.476

alignment block:
US-09-783-320-4 x BC019916 ..

Align seg 1/1 to: BC019916 from: 1 to: 2134

1 MetGIuLyStYrVaLaRgLeuGlnLysIlEgLyGluGlySerPheGly 17
104 ATGGATGACATACATGCTCTGAGATGATGGGAGGGCTCTCGGACG 153
17 sAlaIleuValLysSerThrGluAspGlyArgGlnTyValIlElysc 34
154 AGCTCTTTGGTTCAGCATGAACAGTAAATCATGTTGCCATGAAG 203
34 LuIleAsnIlEserArgMetSerSerLysGluArgGluGluSerArg 50
204 AAATAGAGCTTCCCAAG...TCTTCTCTAATACACAGAAATCTAGAG 250
51 GluValAlaValIleuAlaAsnMetLysHisProAsnIleValGlnTy 67
251 GAGGCTGTTCTTTAGCCAAATGAACCCCTAATATTGTGCTGCCA 300
67 gLusErPheGluGluAsnGlySerLeuTyRlIleValMetAspTyCys 84
301 AGAATCATTTGACGTGAAGACACTGTATATTGTGATGATGACTGTC 350
84 LuGlyGlyAspLeuPheLysArgIleAsnAlaGlnLysGlyValLeuP 100
351 ATGGAGGGGATCTAATGCAAAAGATTAACACGCAAGAAAGAAAGTA 400
101 GlnGluAspGlnIleuAspTrpPheValGlnIleCysLeuAlaLeu 117
401 CCTGAAGACATGATCTAATATTGTTTACCCTAATGCTGCTGAGATA 450
117 sHisValHisAspArgLysIleLeuHisArgAspIleLysSerGlnAs 134
451 TCACATTCACAAGAAACGTGTGTACACAGAGATATCAAGTCCAAAG 500
134 LePheLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyLe 150
501 TCTTCTCATCTCAGAAATGTAATTTGGAGACTTGGATCTGCC 550

151 ArgValLeuAsnSerThrValGluLeuAlaArgThrCysIlEgLyThrPr 167
551 CGTCTCTCTCCAAATCCGATGCGATTCCTGTACTATGTCGGAATCC 600
167 oTyTyTyLeuSerProGluIleCysGluAsnLysProTyRasAsnLys 184
601 TTATATATGTCCTCCAGAAATTTGGAAACCGCTTATACATATAAA 650
184 eAspIleThrPalalaLeuGlyCysValLeuTyGluLeuCysThrLeu 200
651 GTGACATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 700
201 HisAlaPheGluAlaGlySerMetLysAsnLeuValLeuLysIle 217
701 CATCATTTCAAGCAATATGTTGAAAAATCTTATCTCAAGATATGTC 750
217 rGlySerPheProProValSerLeuHisTySerTyRasLeuArgSer 234
751 AGGTCGATCATGTCACCTGCTCTCATATTCATGAACTTCAGTTC 800
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerVal 250
801 TAGTCAAGCAGATGTTAAAGGAATCCCTCACATCGCCCTCGCTACA 850
251 SerIleLeuGluLysGlyPheIleAlaLysArgIleGluLysPheLe 267
851 ACCGTTCTCTCGAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
267 rProGlnLeuIleAlaGluGluPheCysLeuLysThrPheSerLys 284
901 CCCGAAATCATCATGAA..... 919
284 LysSerGlnProIleProAlaLysArgProAlaSerClyGlnAsnSer 300
919 919
301 SerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyGly 317
919 919
317 eProLeuAlaTyLysLysTyGlyAspLysLysLeuHisGluLysP 334
920TATGTCAGAGAGATTTAGAAAGAAATTA. 949
334 roLeuGlnLysHisLysGlnAlaHisGlnThrProGluLysArgVal 350
950AATCGAAGCATACACACCAAGAAACAAACAAAC 985
351 ThrGlyGluGluArgArgLysIleSerGluGluAlaAlaArgLysArg 367
986 CCCAGCAGATTCAGATAGCTTTGGAAATGAAGACAGACAGTCCAA 1033
367 gLeuGluPheIleGluLysGluLysLysGlnLysAspGlnIleIle 384
1034GAGAAACACAGATGAAGAGGTATACCTATGATT 1070
384 euMetLysAlaGluGlnMetLysArgGlnGluLysLeuArgLeuArg 400
1071 TGGAAACCATTAATGA...AATTTAGTTGAAGTCAATTTGAGAAGA 1114
401 IleAsnArgAlaArgGluGlnGlyTrpArgAsnVal 412
1115 GTAAACAGA...GAGAAAAAGCTAATTAAGTCAATC 1147

seq_name: gb_pr:HUMSTRK2A

seq_documentation_block:

LOCUS HUMSTRK2A 3698 bp mRNA linear PRI 09-AUG-1994
DEFINITION Human protein serine/threonine kinase stk2 mRNA, complete cds.
ACCESSION L20321
VERSION L20321.1 GI:348244
KEYWORDS protein serine/threonine kinase; serine/threonine kinase.


```

14002 ...ATPAGAAAAAGAAATTGGAAGCCCTGTGCGGAGGACGCAAAACT 13956
397 gLeuGluArg.....IleasnArgAlaArgGluG 407
13955 ACTGTGAAAAGCGAGCCAAAGCCGACGACCTGTGGAGAAACAAAGGAG 13906
407 InGlyTrpArgsnValLeuSerAlaGlySerGlyGluValLysAla 423
13905 AGCAATCG..... 13898
424 ProPheLeuGlySerGlyGlyThrIleAlaProSerSerPheSerSer 440
13897 .....GAGCG 13893
440 gGlyGlnIleArgIleHisAlaIlePheAspGlnMetGlnGlnAla 457
13892 CAACATGAGGAGCAT.....GCAGAGGAAGTTAGCGAGGCC 13855
457 rGAlaGluAspAsnGluAlaLysTrpLysArgGluIle..... 469
13854 AAGGAGAGGAGGCTGTGCACGTGACAAAGAGAACAAACGCTCTAAC 13805
470 .....TyrGlyArgGlyLeuProGluArgGlnLysG 480
13804 GATPAGCCGTGATCGATACGGTGAATGCT...GGACGGCGTCTCATTT 13758
480 yGlnLeuAlaValAlaGluArgAlaLysGlnValGluPheLeuGlnArg 497
13757 GAAATGCCACCTTCTCCACAGCAGCAGCAGAGTTTACCCTGAGATGC 13708
497 yArgGluAlaMetGlnAsnLysAlaArgAlaGluGlnHisMetValTyr 513
13707 GCCCGCAGCAGCCCTGAAATAG..... 13685
514 LeuAlaArgLeuArgGlnIleArgLeuGlnAsnPheAsnGluArgGln 530
13684 .....CAGCGCTTATGAGAGAGAGAA 13662
530 nIleLys.....AlaI 534
13661 ACTAGAGAACCACTCTCGCTTACGGGCAACCCACAGCGAATGGGA 13612
534 yLeu.....ArgGlyGlnLysGlnAlaAsnHis...SerGln 546
13611 AATTACCAACCGAATAGGGAGGAGAAAGCGGTGTAACCTCTGGCAAT 13562
547 GlyGlnGlnGlySer.....GlnGluAlaAspMetArgAr 558
13561 GGTCAAGACCGCCACACGATATCACCACTAGACATATGAGACAGCCG 13512
558 glySylSllleGlnSerLeuLys.....AlaHisAlaAsnAlaArgAla 573
13511 CAAGGCTTATTTGGATATGCGGCGGAGCAGAGACAGACGAGAG 13462
573 lAValLeuLysGlu.....GlnLeuGluArgLysArg 583
13461 CACTGCTTATGAGGACAGTCCAAAGCGCTGTGAGTGGAAACAACC... 13415
584 LysGluAlaTyrGluArgGluLysLysValTrpGluGlnHisLeuValAl 600
13414 GAAGCGGCGCCCTTAAGGAATCGCAATG.....CCACACATATGAA 13371
600 aLysGlyVal.....LysSerSer 607
13370 AAAGCTTGACACACAGCAACGCGCGCAAGCAATGATACAAACCAAG 13321
607 sPValSerProPheLeuGlyGlnHisGlnIleGly...GlySer 620
13320 ATGTTTGCTTGTCTATGCAAAAGCGACCGCAACCAATGAAGGCCG 13271
621 ProSerLysGlnGlnMetArgSerValIleSerValThrSerAlaLeu 637
13270 CCGTGCCCAAA.....CTGAGCGCAACACATCCCGCCA 13236
637 sGluValGlyVal.....AspSerSerLeuThrAspThrArgGlnHis 652
13235 TAGGGTTGGAGTCTACCTGAAAGTATGTCGCCACATTCCTCCAGATGTG 13186
652 erGlnMetGlnLysThrAsnAsnAlaIleSerSerLysArgGluIle 668
13185 GGACGAG..... 13178
669 LeuArgArgLeuAsnGlnLysAsnLeuLysAlaGlnGluAspGluLysGly 685
13177 .....GTGAAGAGACGAGGATTAACGCACT 13149
685 tGlnAsnLeuSerAsp.....ThrPheGlnLysAsnValHisGluAsp 700
13148 CCAGACTCTTATGACAGGCGCTGCACAAAGAGGCGCATCTCGAGAGTG 13099
700 lAlysGlnHisGluLysGluLysSerValSerSerAspArgLysLysTrp 716
13098 ATGAAGACTTCATGATGAAGCGTGTGCTGACTGTGATATACAGTTT 13049
717 GluAlaGlyGln...LeuValIleProLeuAspGluLeuThrLeuAs 732
13048 GTGCTGATGAGCAGCACTCGCTTCCGAAAGTCAAGCCACA..... 13004
732 pThrSerPheSerThrThrGluArgHisThrValGlyGluValIleLys 749
13003 .....GATCCATGATGACCGTATC...GAGTCTCTTCCG 12970
749 euGlyProAsnGlySerProArgArgAlaTrpGlyLysSerProThrAsp 765
12969 TTTTCTTGAGGCGAA.....CTTGGCAGTCC..... 12941
766 SerValLeuLysIleLeuGlnAlaGluLeuGlnLeu..... 778
12940 .....AAACTTGGGA...GCTTACCGCAATGATGATATTAG 12903
779 .....GlnThrGluLeuLeuGlnAsnThrThrIleArgSerGlnLys 793
12902 TGCAGAGCAGCAGCAACCAATGACAGCT.....GTTCCGATGAATTC 12859
793 erProGlnGlyGluLysTyrLysProLeuIle 803
12858 CAGAGCCCAACATTAATACATCCCAATCATT 12827

```

THIS PAGE BLANK (uspto)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 15, 2002, 07:03:33 ; Search time 54.14 Seconds
(without alignments)
547,704 Million cell updates/sec

Title: US-09-783-320-4
Perfect score: 6243
Sequence: 1 MEKVRLQKIGSGSGFKAIL.....YAKILHLVADGAYQEDNDE 1214

Scoring table: BIOSIM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents-AA:*
1: /cgn2_6/prodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/prodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/prodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/prodata/1/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/prodata/1/1aa/Backfilltest.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1734	27.8	345	3	US-09-173-581-6
2	1734	27.8	345	4	US-09-420-915-6
3	599.5	7.8	979	4	US-08-870-529-2
4	485.5	7.8	1360	4	US-09-393-569-2
5	485.5	7.8	302	3	US-09-221-235-2
6	485.5	7.8	302	4	US-09-221-928-2
7	485.5	7.8	302	4	US-09-221-527-2
8	485.5	7.8	302	4	US-09-221-236-2
9	485.5	7.8	302	4	US-09-221-418-2
10	485.5	7.8	302	4	US-09-221-245-2
11	485.5	7.8	302	4	US-09-163-115-2
12	485.5	7.8	302	4	US-09-221-528-2
13	485.5	7.8	302	4	US-09-593-553-2
14	485.5	7.8	302	4	US-09-221-237-2
15	485.5	7.8	302	4	US-09-221-237-2
16	451	7.2	233	2	US-08-112-709-1
17	451	7.2	233	3	US-09-111-444-1
18	451	7.2	233	4	US-09-541-228-1
19	451	7.2	233	4	US-08-685-871-1
20	436	7.0	293	3	US-09-049-671-1
21	436	7.0	293	3	US-09-295-068-1
22	432.5	6.9	993	4	US-09-060-410-4
23	430	6.9	647	3	US-09-031-563-7
24	430	6.9	648	3	US-09-031-563-5
25	430	6.9	1315	3	US-09-031-563-2
26	427	6.8	1315	3	US-09-293-505-10
27	426	6.8	1386	2	US-08-685-576-4

28	425	6.8	1388	2	US-08-685-576-1	Sequence 1, Appl
29	412	6.6	260	3	US-09-031-563-24	Sequence 24, Appl
30	407.5	6.5	925	1	US-08-252-995D-4	Sequence 4, Appl
31	407.5	6.5	925	2	US-08-834-108-4	Sequence 4, Appl
32	405	6.5	1001	4	US-09-060-410-2	Sequence 62, Appl
33	397	6.4	900	2	US-08-630-822A-62	Sequence 62, Appl
34	387	6.4	900	2	US-09-005-069-62	Sequence 62, Appl
35	384	6.2	239	2	US-08-712-709-7	Sequence 7, Appl
36	384	6.2	239	3	US-09-049-671-3	Sequence 3, Appl
37	384	6.2	239	3	US-09-295-068-3	Sequence 3, Appl
38	384	6.2	239	4	US-09-111-444-7	Sequence 7, Appl
39	384	6.2	729	4	US-08-541-228-7	Sequence 7, Appl
40	384	6.2	729	4	US-08-677-298-2	Sequence 2, Appl
41	381	6.1	416	1	US-08-252-995D-2	Sequence 2, Appl
42	381	6.1	416	2	US-08-834-108-2	Sequence 2, Appl
43	381	6.1	464	1	US-08-252-995D-6	Sequence 6, Appl
44	381	6.1	464	2	US-08-834-108-6	Sequence 6, Appl
45	380.5	6.1	685	2	US-08-878-989-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-173-581-6
Sequence 6, Application US/09173581A
Patent No. 6013455
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Tang, Y. Tom
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
APPLICANT: Gorgone, Gina
APPLICANT: Azimzai, Yalda
APPLICANT: Lu, Aina
TITLE OR INVENTION: Protein Kinase Homologs
FILE REFERENCE: PR-0614 US
CURRENT APPLICATION NUMBER: US/09/173,581A
CURRENT FILING DATE: 1998-10-15
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PERL Program
SEQ ID NO 6
LENGTH: 345
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE: -
OTHER INFORMATION: 1567782
US-09-173-581-6

Query Match	27.8%	Score 1734;	DB 3;	Length 345;
Best Local Similarity	90.5%	Pred. No. 1.3e-89;		
Matches 341;	Conservative	3;	Mismatches 1;	Indels 32; Gaps 1;
QY	1	MEKVRLQKIGSGSGFKAILVKSTEDGROYVIREINISRMSSKEREESREYAVLANKH	60	
DB	1	MEKVRLQKIGSGSGFKAILVKSTEDGROYVIREINISRMSSKEREESREYAVLANKH	60	
QY	61	PNIVYORRESFENGSLYIVMDYCEGDLFFKRNQKGVLFQEDQIIDWFOVQICLAKHVN	120	
DB	61	PNIVYORRESFENGSLYIVMDYCEGDLFFKRNQKGVLFQEDQIIDWFOVQICLAKHVN	120	
QY	121	DKTILHROIKSQNIFLTGDTGVQGLGFGIARVUNSVTELARTCIGPYLSPICENKPY	180	
DB	121	DKTILHROIKSQNIFLTGDTGVQGLGFGIARVUNSVTELARTCIGPYLSPICENKPY	180	
QY	89	DKTILHROIKSQNIFLTGDTGVQGLGFGIARVUNSVTELARTCIGPYLSPICENKPY	148	
DB	89	DKTILHROIKSQNIFLTGDTGVQGLGFGIARVUNSVTELARTCIGPYLSPICENKPY	148	
QY	181	NKSDIWAIGCVLYELCTLKAHAFAGSMKNLYLTIISGSPVSLHYSVDLSVLSQLEK	240	
DB	181	NKSDIWAIGCVLYELCTLKAHAFAGSMKNLYLTIISGSPVSLHYSVDLSVLSQLEK	240	
QY	149	NNKSDIWAIGCVLYELCTLKAHAFAGSMKNLYLTIISGSPVSLHYSVDLSVLSQLEK	208	
DB	149	NNKSDIWAIGCVLYELCTLKAHAFAGSMKNLYLTIISGSPVSLHYSVDLSVLSQLEK	208	
QY	241	RNRDRPSVNSILIEKGFIAKRIEKLTSQILAEFCLKTSKSGQPIAKRPAASQNSI	300	
DB	241	RNRDRPSVNSILIEKGFIAKRIEKLTSQILAEFCLKTSKSGQPIAKRPAASQNSI	300	

Wed May 15 14:13:55 2002

us-09-783-320-4.ra1

Page 2

Db 209 RNPDRPSVNSILEKGFIAKRIEFLSPOLIAEFCLKTFKSGSOPIPAKRPASGNSI 268
301 SVMPAOKITTKPAAYGIPLAYKKYGGDKLHEKKRPLOKHQAHOHPKRVNTGERRKISE 360
269 SVMPAOKITTKPAAYGIPLAYKKYGGDKLHEKKRPLOKHQAHOHPKRVNTGERRKISE 328
Qy 361 EAARKRRLEFIEKERRK 377
Db 329 EAARKRRLEFIEKERRK 345

RESULT 2
US-09-420-915-6
Sequence 6, Application US/09420915
Patent No. 626497
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Tang, Y. Tom
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
APPLICANT: Gorgone, Gina
APPLICANT: Azimzal, Yalda
APPLICANT: Lu, Alina
TITLE OF INVENTION: Protein Kinase Homologs
FILE REFERENCE: PF-0614 US
CURRENT APPLICATION NUMBER: US/09/420,915
CURRENT FILING DATE: 1999-10-20
EARLIER APPLICATION NUMBER: US 09/173,581
EARLIER FILING DATE: 1998-10-15
NUMBER OF SEQ. ID NOS: 18
SOFTWARE: PERL Program
SEQ. ID NO. 6
LENGTH: 345
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE: -
OTHER INFORMATION: 1567782
US-09-420-915-6

Query Match 27.8%; Score 1734; DB 4; Length 345;
Best Local Similarity 90.5%; Pred. No. 1.3e-89;
Matches 341; Conservative 3; Mismatches 1; Indels 32; Gaps 1;
Qy 1 MEKYRLQKIGESFGKAILVSTEDGROYVKEINISRMSSKEREESRREVAVALANKMH 60
Db 1 MEKYRLQKIGESFGKAILVSTEDGROYVKEINISRMSSKEREESRREVAVALANKMH 60
Qy 61 PIVVOYRESFENGSLVYIMDYCGDLPFRINAQGVLPQEDQIIDMVOICLAKKHVH 120
Db 61 PIVVOYRESFENGSLVYIMDYCGDLPFRINAQGVLPQEDQIIDMVOICLAKKHVH 120
Qy 121 DRKILHRIRKSONIFLRKDGTVOLGDFGIAVNSTVELARTGIPYVLSPEICENKRY 180
Db 89 DRKILHRIRKSONIFLRKDGTVOLGDFGIAVNSTVELARTGIPYVLSPEICENKRY 148
Qy 181 NKSSTIALGCVLYELCTLKAHFAFGSMKMLVLIISGSPVPSLHSTDLNSVQLFK 240
Db 149 NKSSTIALGCVLYELCTLKAHFAFGSMKMLVLIISGSPVPSLHSTDLNSVQLFK 208
Qy 241 RNPDRPSVNSILEKGFIAKRIEFLSPOLIAEFCLKTFKSGSOPIPAKRPASGNSI 300
Db 209 RNPDRPSVNSILEKGFIAKRIEFLSPOLIAEFCLKTFKSGSOPIPAKRPASGNSI 268
Qy 301 SVMPAOKITTKPAAYGIPLAYKKYGGDKLHEKKRPLOKHQAHOHPKRVNTGERRKISE 360
Db 269 SVMPAOKITTKPAAYGIPLAYKKYGGDKLHEKKRPLOKHQAHOHPKRVNTGERRKISE 328
Qy 361 EAARKRRLEFIEKERRK 377
Db 329 EAARKRRLEFIEKERRK 345

Db 329 EAARKRRLEFIEKERRK 345

RESULT 3
US-08-870-529-2
Sequence 2, Application us/08870529
Patent No. 6080557
GENERAL INFORMATION:
APPLICANT: Sims, John E.
APPLICANT: Virca, G. Duke
APPLICANT: Baird, Timothy A.
APPLICANT: Anderson, Dik M.
TITLE OF INVENTION: IL-1/TNF-(ACTIVATED KINASE (ITAK)
TITLE OF INVENTION: AND METHODS OF MAKING AND USING THE SAME
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/870,529
FILING DATE: 06-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 480052.418
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 979 amino acids
TYPE: amino acid
TOPOLOGY: 1linear
MOLECULE TYPE: protein
US-08-870-529-2

Query Match 9.6%; Score 599.5; DB 3; Length 979;
Best Local Similarity 37.5%; Pred. No. 7.4e-26;
Matches 124; Conservative 71; Mismatches 113; Indels 23; Gaps 5;
Qy 4 YVRLOKIGESFGKAILVSTEDGROYVKEINISRMSSKEREESRREVAVALANKMHPT 63
Db 52 YIRIVYGRARPEALYRTEDDSLVYWKEDVLRLEKERRALNIVIALLOHNDI 111
Qy 64 VOYRESFENGSLVYIMDYCGDLPFRINAQGVLPQEDQIIDMVOICLAKKHVH 123
Db 112 IAYVNHMDVTLLIELEKNGENLDRKLNQKMLFEENVMVLYQIYSVSCIHKAG 171
Qy 124 ILHRDIKSONIFLRKDGTVOLGDFGIAVNSTVELARTGIPYVLSPEICENKRY 183
Db 172 ILHRDIKSONIFLRKDGTVOLGDFGIAVNSTVELARTGIPYVLSPEICENKRY 231
Qy 184 SDIWMALGCVLYELCTLKAHFAFGSMKMLVLIISGSPVPSLHSTDLNSVQLFK 241
Db 232 SDIWMALGCVLYELCTLKAHFAFGSMKMLVLIISGSPVPSLHSTDLNSVQLFK 208
Qy 242 RNPDRPSVNSILEKGFIAKRIEFLSPOLIAEFCLKTFKSGSOPIPAKRPASGNSI 301
Db 292 RNPDRPSVNSILEKGFIAKRIEFLSPOLIAEFCLKTFKSGSOPIPAKRPASGNSI 268
Qy 302 VMPAOKITTKPAAYGIPLAYKKYGGDKLHEKKRPLOKHQAHOHPKRVNTGERRKISE 360
Db 292 VMPAOKITTKPAAYGIPLAYKKYGGDKLHEKKRPLOKHQAHOHPKRVNTGERRKISE 328

Wed May 15 14:13:55 2002

us-09-783-320-4.rai

Page 4

Matches 102; Conservative 48; Mismatches 101; Indels 11; Gaps 6;

[illegible]

RESULT 6
US-09-221-928-2

Sequence 2. Application US/09221928
Patient No. 6121030
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: MNT-030
CURRENT APPLICATION NUMBER: US/09/221,928
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE:
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 302
TYPE: PRT
ORGANISM: Homo sapiens
US-09-221-928-2

Query Match 7.88; Score 485.5; DB 3; Length 302;

Best Local Similarity	38.98;	Pred. No. 4e-20;
Matches 102;	Conservative 48;	Mismatches 101;
	Indels 11;	Gaps 6

Qy	8	OKIGESPEKALUWSTEDGNOYIKETENI--SWSSEKESEBESREPAVALAMKPNPIVOY	66
Db	38	KKIGROFSEVYRAACILDDGPVALKKVOYFJDLMDAKARADCIKEITDLKOLNHPVITY	97
Qy	67	RESEFENGSLYIMDYCEGDL---PRINAGCVLFQEDQILDMFVOICLAKHYDRK	123
Db	98	YASFIENENINTYLEADJGDSPKRIHFEKKÖR--LIPETVWKIFVQCSLEHMSRR	156
Qy	124	ILHRDIKSONIEFLTDGVYQJADFGIARVLNSTVEIARPCIGTPYILSPJECKPNYNNK	183
Db	157	VMRHDIKPNWVFLTATGVYKJADJGLGRFESSKTTAAHSLVSGTPYMSBERIHENGNNK	216
Qy	184	SDIUALGCVLYELCTLTKHAFAENGSK--NVLAKITIGSFPVY-SLHYSYDLKSLVSOLF	240
Db	217	SDIWSIGCLLYEAAALQSPFYGDKNMNYLSLCKIIEQCDYPLPSDHYSEELROLVNNCIN	276
Qy	241	RNPDRPSVNSILKGFIAKRI	262
Db	277	PDEKRPDYTVYD--VAKRH	295

RESULT 7
US-09-221-527-2
; Sequence 2, Application US/09221527
; Patent No. 6146832

```

GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: MNT-050
CURRENT APPLICATION NUMBER: US/09/221,527
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE:
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 302
TYPE: PRT
ORGANISM: Homo sapiens
US-09-221-527-2

```

Query Match 7.88; Score 485.5; DB 4; Length 302;

Best Local Similarity	38.98;	Pred. No. 4e-20;	
Matches 102; Conservative	48;	Mismatches 101;	Indels 11; Gaps 6

QY 8 QKIGEGSGKALVKSTEDGRÖYVIKEINI-SRMSSKEREESRREVAVLANKHPNIVQY 66

Db 38 K K I G R G Q F S E V Y R A C C L D G V P A L K K V Q I F D L M D A K A R A D C I K E I D L K Q L N H P N V I K Y 97

```
QY      67 RESFEENGSLIVMDYCEGDL--FKRINAQKGVLFQEDQILDWFVQICLALKHVHDK 12:
```

DB 98 YASF IEDNEENI VEEADAVGUSKMIANHE ANVAK - LIFEKI VMA IF VQUCSAEENMHSK 130

[illegible][illegible]

Db 217 SDIWSIGCIIYEMAIOSPFYGDKNILYSLCKITEOCDDYPLPSDHYSEELBOLYNMCIN 276

QY 241 RNPRDRPSVNSILEKGFIAKRI 262

Db 277 PDPEKRPDVTYYD---VAKRM 295

RESULT 8

US-09-221-236-2

Sequence 2, Application US/09221236

Patent No. 6146841
GENERAL INFORMATION:

APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
PRIOR REFERENCE: INT-050

;; CURRENT APPLICATION NUMBER: US/09/221,236
: CURRENT FILING DATE: 1998-12-28

EARLIER APPLICATION NUMBER: 09/165,113
EARLIER FILING DATE: 1998-09-29

SOFTWARE: PatentIn Ver. 2.0

LENGTH: 302

ORGANISM: Homo sapiens
; ITS-09-221-236-2

Query Match

Matches	102;	Conservative	48;	Mismatches	101;	Indels	11;	Gaps	6
---------	------	--------------	-----	------------	------	--------	-----	------	---

QY 8 QKIGESFGKALVKSTEDGRÖYVIKEINI-SRMSKEREESRREVAVLANKHPNIVÖY 666

Db 38 KRIIGQFSEVYRACLDGPVALRKVQLFLDLMDAKARADCIRKELDLKQLNHPNVIKI 9/

```
0 / RESFEENGSLIWMDCGGDL--FARLNAGNGVLEQEDQLDWFVGICLGNHVNHNK
|||: |||: : ||| | | : |||:| ||:| |:
```

```

; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-245-2

Query Match          7.8%; Score 485.5; DB 4; Length 302;
Best Local Similarity 38.9%; Pred. No. 4e-20;
Matches 102; Conservative 48; Mismatches 101; Indels 11; Gaps 6;

OY 8 OKIGESFGKALIVSTEDGROYVKEINT--SRMSKERESESRREAVAVLANKHNPVIOY 66
DB 38 KTIIGROFSEVYRAACLDGVPVALKVOIFDLMDAKARADCIKEIDLKOLNHPNVIXY 97
OY 67 RESFENGSLYIVMDYCEGGDL--FKRINOKGVLFQEDQILDMFVOICLAKVHNDK 123
DB 98 YASFIEDNELNIVLEADAGDLSRMKHKOKR-LIPERTVWKYFVQCSALEHMSRR 156
OY 124 ILHRDIKSONIFLTGDTVOLGDFGIARVLSNIVELARTCIGTPYVLSPEICENKPYNNK 183
DB 157 VMHRDIKIPANVFITATGVVKKLDLGLGRFSSKTTAAHSLVGTPYMSPERIHENGYNFK 216
OY 184 SDIWAIGCVLYELCTLKAHFEAGSMK--NLVKIISGSPPV-SLHYSYDLRSLSQLEK 240
DB 217 SDIWSIGCLLYEMALQSPFYGDKNMLYSLCKRIEODCDYPLPSDHSSELRQLVNMCIN 276
OY 241 RNPDRPSVNSILEKGFIAKRI 262
DB 277 PDPEKRDPVTVYVD--VAKRM 295

; RESULT 9
US-09-221-416-2
; Sequence 2, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNT-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-416-2

Query Match          7.8%; Score 485.5; DB 4; Length 302;
Best Local Similarity 38.9%; Pred. No. 4e-20;
Matches 102; Conservative 48; Mismatches 101; Indels 11; Gaps 6;

OY 8 OKIGESFGKALIVSTEDGROYVKEINT--SRMSKERESESRREAVAVLANKHNPVIOY 66
DB 38 KTIIGROFSEVYRAACLDGVPVALKVOIFDLMDAKARADCIKEIDLKOLNHPNVIXY 97
OY 67 RESFENGSLYIVMDYCEGGDL--FKRINOKGVLFQEDQILDMFVOICLAKVHNDK 123
DB 98 YASFIEDNELNIVLEADAGDLSRMKHKOKR-LIPERTVWKYFVQCSALEHMSRR 156
OY 124 ILHRDIKSONIFLTGDTVOLGDFGIARVLSNIVELARTCIGTPYVLSPEICENKPYNNK 183
DB 157 VMHRDIKIPANVFITATGVVKKLDLGLGRFSSKTTAAHSLVGTPYMSPERIHENGYNFK 216
OY 184 SDIWAIGCVLYELCTLKAHFEAGSMK--NLVKIISGSPPV-SLHYSYDLRSLSQLEK 240
DB 217 SDIWSIGCLLYEMALQSPFYGDKNMLYSLCKRIEODCDYPLPSDHSSELRQLVNMCIN 276
OY 241 RNPDRPSVNSILEKGFIAKRI 262
DB 277 PDPEKRDPVTVYVD--VAKRM 295

; RESULT 10
US-09-221-245-2
; Sequence 2, Application US/09221245
; Patent No. 6180358
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNT-050
; CURRENT APPLICATION NUMBER: US/09/221,245
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: US 09/163,115
; EARLIER FILING DATE: 1998-09-29

```

```

; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-245-2

Query Match          7.8%; Score 485.5; DB 4; Length 302;
Best Local Similarity 38.9%; Pred. No. 4e-20;
Matches 102; Conservative 48; Mismatches 101; Indels 11; Gaps 6;

OY 8 OKIGESFGKALIVSTEDGROYVKEINT--SRMSKERESESRREAVAVLANKHNPVIOY 66
DB 38 KTIIGROFSEVYRAACLDGVPVALKVOIFDLMDAKARADCIKEIDLKOLNHPNVIXY 97
OY 67 RESFENGSLYIVMDYCEGGDL--FKRINOKGVLFQEDQILDMFVOICLAKVHNDK 123
DB 98 YASFIEDNELNIVLEADAGDLSRMKHKOKR-LIPERTVWKYFVQCSALEHMSRR 156
OY 124 ILHRDIKSONIFLTGDTVOLGDFGIARVLSNIVELARTCIGTPYVLSPEICENKPYNNK 183
DB 157 VMHRDIKIPANVFITATGVVKKLDLGLGRFSSKTTAAHSLVGTPYMSPERIHENGYNFK 216
OY 184 SDIWAIGCVLYELCTLKAHFEAGSMK--NLVKIISGSPPV-SLHYSYDLRSLSQLEK 240
DB 217 SDIWSIGCLLYEMALQSPFYGDKNMLYSLCKRIEODCDYPLPSDHSSELRQLVNMCIN 276

; RESULT 11
US-09-163-115-2
; Sequence 2, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNT-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-163-115-2

Query Match          7.8%; Score 485.5; DB 4; Length 302;
Best Local Similarity 38.9%; Pred. No. 4e-20;
Matches 102; Conservative 48; Mismatches 101; Indels 11; Gaps 6;

OY 8 OKIGESFGKALIVSTEDGROYVKEINT--SRMSKERESESRREAVAVLANKHNPVIOY 66
DB 38 KTIIGROFSEVYRAACLDGVPVALKVOIFDLMDAKARADCIKEIDLKOLNHPNVIXY 97
OY 67 RESFENGSLYIVMDYCEGGDL--FKRINOKGVLFQEDQILDMFVOICLAKVHNDK 123
DB 98 YASFIEDNELNIVLEADAGDLSRMKHKOKR-LIPERTVWKYFVQCSALEHMSRR 156
OY 124 ILHRDIKSONIFLTGDTVOLGDFGIARVLSNIVELARTCIGTPYVLSPEICENKPYNNK 183
DB 157 VMHRDIKIPANVFITATGVVKKLDLGLGRFSSKTTAAHSLVGTPYMSPERIHENGYNFK 216
OY 184 SDIWAIGCVLYELCTLKAHFEAGSMK--NLVKIISGSPPV-SLHYSYDLRSLSQLEK 240
DB 217 SDIWSIGCLLYEMALQSPFYGDKNMLYSLCKRIEODCDYPLPSDHSSELRQLVNMCIN 276

```



```
; Patent No. 5863780
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,709
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 233 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
; US-08-712-709-1
```

```
Query Match 7.28; Score 451; DB 2; Length 233;
Best Local Similarity 38.98; Pred. NO. 2.4e-18;
Matches 91; Conservative 48; Mismatches 85; Indels 10; Gaps 5;

OY 40 MSKREESRRREAVLANMKHPNIVQRESFEENGSLIYMDYCEGDL--FKRINAK 96
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2 MAAKAKODCVKEIGLKQINHPNIIKYIDSTEDNELNIVLEADAGDLPOMIKFKKK 61
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 97 GVLFQEDQILDWFVQICLALKVHDRLIHRDIKSONIFLTKDGVOLGDFGIARYLNT 156
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 62 R-LIPERTVWKYFVQLCSAVEHMSRRVMHRDIPANVFTATGVVKGDLGGRFSS 120
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 157 VELARFCIGTPYYLSPEICENKPYNNKSDIWLGCVLIELCTLKHAFEAGSMK--NLVIK 214
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 121 TTAHSLVCGTPYYMSPERHENGYNFKSDIMSLGCLLYEMALQSPYGDKNLFSLCOK 180
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 215 IIGSGPPV-SLHYSYDLRSIVSOLFKNRPDRPSVNSILEKGFIAKRIEKL 267
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 181 IEQCDYPLPGEHYSEKLELYSMCICPDHPQRPDIGXVH--VAKOMHIWMS 231
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: May 15, 2002, 07:53:47
Job time: 3014 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 15, 2002, 07:53:00 ; Search time 262.69 Seconds
(without alignments)

1626.649 Million cell updates/sec

Title: US-09-783-320-4
Perfect score: 6243
Sequence: 1 MEKVRLQKIGESFGKAIL.....YAKILHLVADGAYOEINDE 1214

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%
Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

- 1: /cgn2_6/ptodata/1/paa/PCUS.COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06.COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07.COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US08.COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US081.COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US082.COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US083.COMB.pep.*
- 8: /cgn2_6/ptodata/1/paa/US084.COMB.pep.*
- 9: /cgn2_6/ptodata/1/paa/US085.COMB.pep.*
- 10: /cgn2_6/ptodata/1/paa/US086.COMB.pep.*
- 11: /cgn2_6/ptodata/1/paa/US087.COMB.pep.*
- 12: /cgn2_6/ptodata/1/paa/US088.COMB.pep.*
- 13: /cgn2_6/ptodata/1/paa/US089.COMB.pep.*
- 14: /cgn2_6/ptodata/1/paa/US090.COMB.pep.*
- 15: /cgn2_6/ptodata/1/paa/US091.COMB.pep.*
- 16: /cgn2_6/ptodata/1/paa/US092.COMB.pep.*
- 17: /cgn2_6/ptodata/1/paa/US093.COMB.pep.*
- 18: /cgn2_6/ptodata/1/paa/US094.COMB.pep.*
- 19: /cgn2_6/ptodata/1/paa/US095.COMB.pep.*
- 20: /cgn2_6/ptodata/1/paa/US096.COMB.pep.*
- 21: /cgn2_6/ptodata/1/paa/US097.COMB.pep.*
- 22: /cgn2_6/ptodata/1/paa/US098.COMB.pep.*
- 23: /cgn2_6/ptodata/1/paa/US099.COMB.pep.*
- 24: /cgn2_6/ptodata/1/paa/US100.COMB.pep.*
- 25: /cgn2_6/ptodata/1/paa/US101.COMB.pep.*
- 26: /cgn2_6/ptodata/1/paa/US060.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6243	100.0	1214	21	US-09-783-320-4
2	6237	99.9	1214	18	US-09-488-725A-2356
3	6213	99.5	1242	18	US-09-488-725A-2355
4	5162	82.7	1007	21	US-09-783-320-6
5	5138	82.3	1035	21	US-09-783-320-2
6	4735.5	75.9	966	18	US-09-488-725A-5927
7	4735.5	75.9	966	18	US-09-488-725A-5928

8	3787	60.7	778	26	US-60-299-426-30	Sequence 30, Appl
9	2785	44.6	616	22	US-09-824-583-4	Sequence 4, Appl
10	2785	44.6	616	22	US-09-873-404-4	Sequence 4, Appl
11	1734	27.8	345	22	US-09-870-962-6	Sequence 6, Appl
12	1283	20.6	254	21	US-09-715-427-35	Sequence 35, Appl
13	1283	20.6	254	22	US-09-898-837A-35	Sequence 35, Appl
14	1146	18.4	256	21	US-09-760-446A-1590	Sequence 1590, Ap
15	1060	17.0	640	22	US-09-824-583-2	Sequence 2, Appl
16	1057	16.9	640	22	US-09-824-583-2	Sequence 2, Appl
17	1057	16.9	640	22	US-09-824-583-2	Sequence 2, Appl
18	1057	16.9	640	22	US-09-824-583-2	Sequence 2, Appl
19	882	14.1	220	21	US-09-760-446A-1590	Sequence 2215, Ap
20	875	14.0	506	1	PCT-US01-04098A-1006	Sequence 1006, Ap
21	848	13.6	836	26	US-60-208-965-221	Sequence 2974, Ap
22	847.5	13.6	836	26	US-60-213-178-724	Sequence 724, Ap
23	847.5	13.6	841	19	US-09-538-092-1183	Sequence 1183, Ap
24	841.5	13.5	269	26	US-60-200-366-127	Sequence 127, Ap
25	841.5	13.5	269	26	US-60-248-798-347	Sequence 347, Ap
26	806.5	12.9	510	26	US-60-106-889-6	Sequence 6, Appl
27	790.5	12.7	459	19	US-09-538-092-1182	Sequence 1182, Ap
28	786	12.6	458	26	US-09-771-161A-198	Sequence 198, Ap
29	786	12.6	458	26	US-60-200-366-128	Sequence 128, Ap
30	780.5	12.5	416	21	US-09-731-231-6	Sequence 6, Appl
31	758.5	12.1	264	26	US-60-206-600-158	Sequence 158, Ap
32	756.5	12.1	399	26	US-60-213-178-776	Sequence 776, Ap
33	747	12.0	143	1	PCT-US00-26376-111	Sequence 111, Ap
34	743.5	11.9	260	21	US-09-731-231-5	Sequence 5, Appl
35	733.5	11.7	255	21	US-09-715-427-36	Sequence 36, Appl
36	733.5	11.7	255	21	US-09-898-837A-36	Sequence 36, Appl
37	678.5	10.9	234	26	US-60-248-798-113	Sequence 113, Ap
38	622.5	10.0	698	23	US-09-995-530-15	Sequence 15, Appl
39	621.5	10.0	692	23	US-09-992-481-2	Sequence 2, Appl
40	620.5	9.9	692	23	US-09-910-150-5	Sequence 5, Appl
41	604	9.7	645	21	US-09-731-231-2	Sequence 2, Appl
42	599.5	9.6	645	26	US-60-220-038-6	Sequence 6, Appl
43	599.5	9.6	979	10	US-08-633-414-2	Sequence 2, Appl
44	599.5	9.6	979	19	US-09-544-794-2	Sequence 2, Appl
45	599.5	9.6	979	22	US-09-884-001-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-783-320-4
; Sequence 4, Application US/09783320
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hu, Yi
; APPLICANT: Nepomichy, Boris
; APPLICANT: Turner, C. Alexander JR
; APPLICANT: Zambrowicz, Brian
; TITLE OF INVENTION: Novel Human Kinases and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0137-USA
; CURRENT FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: US 60/183,582
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/184,014
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1214
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-783-320-4

Query Match 100.0%; Score 6243; DB 21; Length 1214;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MEKYVRLQKIGESFGKALIVKSTEDGRQYVYKEINISRMSSKREESREAVLANMKH 60
DB 1 MEKYVRLQKIGESFGKALIVKSTEDGRQYVYKEINISRMSSKREESREAVLANMKH 60
QY 61 PNIVQYRESEFENGSLIYIMDCCEGGDLFKRINAKGVLPQEDQILDMFVOICLAKH 120
DB 61 PNIVQYRESEFENGSLIYIMDCCEGGDLFKRINAKGVLPQEDQILDMFVOICLAKH 120
QY 121 DRKILHRDIKSONIFLTQKGTVOIGDFGIARVLNSTVELATFCIGTGYIISPEICENKPY 180
DB 121 DRKILHRDIKSONIFLTQKGTVOIGDFGIARVLNSTVELATFCIGTGYIISPEICENKPY 180
QY 181 NNSSDIHALGCVLYELCTLKHAFEGAGSMKNLVKTIISGSPFVSLHYSDLRSLVSQLFK 240
DB 181 NNSSDIHALGCVLYELCTLKHAFEGAGSMKNLVKTIISGSPFVSLHYSDLRSLVSQLFK 240
QY 241 RNRDRPSVNSIIEKGFIAKRIEKLSPOLIAEECLTFESKFSQPIPAKRPASGONSI 300
DB 241 RNRDRPSVNSIIEKGFIAKRIEKLSPOLIAEECLTFESKFSQPIPAKRPASGONSI 300
QY 301 SVMPAQKITKPAKYGIPLAYKKYGDKLHEKKPLQKHQAHOPEKRVNTGEERKISE 360
DB 301 SVMPAQKITKPAKYGIPLAYKKYGDKLHEKKPLQKHQAHOPEKRVNTGEERKISE 360
QY 361 EAARKRLFEIEKEKQKQDOISLMKAEOMKROEKERLERIRAREQGMNVLSAGSGE 420
DB 361 EAARKRLFEIEKEKQKQDOISLMKAEOMKROEKERLERIRAREQGMNVLSAGSGE 420
QY 421 VKAPFLSGGTIAPSSFSRGQYEHYALFTDOMOOARADENAKKREIYGRGLPERQK 480
DB 421 VKAPFLSGGTIAPSSFSRGQYEHYALFTDOMOOARADENAKKREIYGRGLPERQK 480
QY 481 QLAVERAKOYEELQKREAMONKARABEGHMYLARLRQIRLONNEGOIKAKIRGSKK 540
DB 481 QLAVERAKOYEELQKREAMONKARABEGHMYLARLRQIRLONNEGOIKAKIRGSKK 540
QY 541 EAHNSEOGESEADMRKRIEISIKHANARAALVKEOLEERKRAYREKKVMEHLYA 600
DB 541 EAHNSEOGESEADMRKRIEISIKHANARAALVKEOLEERKRAYREKKVMEHLYA 600
QY 601 KGVKSSVSPPLQGHETGSSPSQOQRSVIYSTSAKKEGVSSLTIDRETSEEKQKNN 660
DB 601 KGVKSSVSPPLQGHETGSSPSQOQRSVIYSTSAKKEGVSSLTIDRETSEEKQKNN 660
QY 661 AISSKREILRLNENLKAODEKGMONLSDFEINVHEDAKHEKESVSSDRKKWEAG 720
DB 661 AISSKREILRLNENLKAODEKGMONLSDFEINVHEDAKHEKESVSSDRKKWEAG 720
QY 721 QLVIPDLDELTLDTSFSTTERHTVGEVYIKLGPNGSPRAWGKSPTSVYKILGAEALQOT 780
DB 721 QLVIPDLDELTLDTSFSTTERHTVGEVYIKLGPNGSPRAWGKSPTSVYKILGAEALQOT 780
QY 781 ELLENTTIRSEISPEGEKKYPLITGEEKVQOCSHSEINPAIVDSVEIKSPFESASQOM 840
DB 781 ELLENTTIRSEISPEGEKKYPLITGEEKVQOCSHSEINPAIVDSVEIKSPFESASQOM 840
QY 841 SLKLEGLLEPDDLETEIIOEPSGTNKKDSLCJTITDWAISEKRTKTSQASDRITIOEN 900
DB 841 SLKLEGLLEPDDLETEIIOEPSGTNKKDSLCJTITDWAISEKRTKTSQASDRITIOEN 900
QY 901 EYSEEDGVSSITVDOLSDIHIEPGTNDOSHKCOVDKSVQPEPFKRVVHSEHLNLPVOVS 960
DB 901 EYSEEDGVSSITVDOLSDIHIEPGTNDOSHKCOVDKSVQPEPFKRVVHSEHLNLPVOVS 960
QY 961 VOCSPESFARSHSLPRKNNKNSLLGLSTGLFDANNPKMLRTCSLPDLSKLFRTLM 1020
DB 961 VOCSPESFARSHSLPRKNNKNSLLGLSTGLFDANNPKMLRTCSLPDLSKLFRTLM 1020
QY 1021 DVPTVGVRODNLDEIDENIKESGSDSEDIVEEETDTDQEOQASMEQLRPOPEE 1080
DB 1021 DVPTVGVRODNLDEIDENIKESGSDSEDIVEEETDTDQEOQASMEQLRPOPEE 1080
QY 1081 YSEEEESVLKNSDVEPTANGTDVAEDBDNPSBSALNEMWHSNDGELIASCECDSVFN 1140

```

```

DB 1081 YSEEEESVLKNSDVEPTANGTDVAEDBDNPSBSALNEMWHSNDGELIASCECDSVFN 1140
QY 1141 HEEELRLHLEOMGFEKFEVEYERIKALTHEDEDENEIECSKIYONILGNHQLYAKILH 1200
DB 1141 HEEELRLHLEOMGFEKFEVEYERIKALTHEDEDENEIECSKIYONILGNHQLYAKILH 1200
QY 1201 LVMAOGAYOEDNDE 1214
DB 1201 LVMAOGAYOEDNDE 1214

RESULT 2
US-09-488-725A-2356
Sequence 2356, Application US/09488725A
GENERAL INFORMATION:
APPLICANT: Hysq Inc
TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
FILE REFERENCE: 784FLPCT
CURRENT APPLICATION NUMBER: US/09/488, 725A
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US09/552, 317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US09/598, 042
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: US09/620, 312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: US09/653, 450
PRIOR FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: US09/662, 191
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: US09/693, 036
PRIOR FILING DATE: 2000-10-19
PRIOR APPLICATION NUMBER: US09/727, 344
PRIOR FILING DATE: 2000-11-29
NUMBER OF SEQ ID NOS: 7144
SOFTWARE: PL_FL_genes_1.0
SEQ ID NO 2356
LENGTH: 1215
TYPE: PRT
ORGANISM: Homo sapiens
US-09-488-725A-2356

Query Match 99.9%; Score 6237; DB 18; Length 1214;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1213; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEKYVRLQKIGESFGKALIVKSTEDGRQYVYKEINISRMSSKREESREAVLANMKH 60
DB 1 MEKYVRLQKIGESFGKALIVKSTEDGRQYVYKEINISRMSSKREESREAVLANMKH 60
QY 61 PNIVQYRESEFENGSLIYIMDCCEGGDLFKRINAKGVLPQEDQILDMFVOICLAKH 120
DB 61 PNIVQYRESEFENGSLIYIMDCCEGGDLFKRINAKGVLPQEDQILDMFVOICLAKH 120
QY 121 DRKILHRDIKSONIFLTQKGTVOIGDFGIARVLNSTVELATFCIGTGYIISPEICENKPY 180
DB 121 DRKILHRDIKSONIFLTQKGTVOIGDFGIARVLNSTVELATFCIGTGYIISPEICENKPY 180
QY 181 NNSSDIHALGCVLYELCTLKHAFEGAGSMKNLVKTIISGSPFVSLHYSDLRSLVSQLFK 240
DB 181 NNSSDIHALGCVLYELCTLKHAFEGAGSMKNLVKTIISGSPFVSLHYSDLRSLVSQLFK 240
QY 241 RNRDRPSVNSIIEKGFIAKRIEKLSPOLIAEECLTFESKFSQPIPAKRPASGONSI 300
DB 241 RNRDRPSVNSIIEKGFIAKRIEKLSPOLIAEECLTFESKFSQPIPAKRPASGONSI 300
QY 301 SVMPAQKITKPAKYGIPLAYKKYGDKLHEKKPLQKHQAHOPEKRVNTGEERKISE 360
DB 301 SVMPAQKITKPAKYGIPLAYKKYGDKLHEKKPLQKHQAHOPEKRVNTGEERKISE 360

```

```

QY 361 EAARKRRLEFTEKROKODIISLMKAEOMKROEKERLERINAROGMRNVLASGSGE 420
|
|
|
Db 361 EAARKRRLEFTEKROKODIISLMKAEOMKROEKERLERINAROGMRNVLASGSGE 420
QY 421 VKAPFLGSGGTAPSSFSRGOYEYHAIFDOMOORADNEAKMKREIYGRGLPEROKG 480
|
|
|
Db 421 VKAPFLGSGGTAPSSFSRGOYEYHAIFDOMOORADNEAKMKREIYGRGLPEROKG 480
QY 481 QLAVERAKOVEEFLORRKEAMONKARAGHMYLARLROIRLONFNEROOIKAKLNGEK 540
|
|
|
Db 481 QLAVERAKOVEEFLORRKEAMONKARAGHMYLARLROIRLONFNEROOIKAKLNGEK 540
QY 541 EANHSGOGSEADMRKRIESTLKAHNAARAVALKEOLEKREKREAVEREKRWEEHLVA 600
|
|
|
Db 541 EANHSGOGSEADMRKRIESTLKAHNAARAVALKEOLEKREKREAVEREKRWEEHLVA 600
QY 601 KGVKSSDVPSPLOHETGSGSPKQOMRSVIVTSALKRVGVSLSLTDTRETSEMOKTN 660
|
|
|
Db 601 KGVKSSDVPSPLOHETGSGSPKQOMRSVIVTSALKRVGVSLSLTDTRETSEMOKTN 660
QY 661 AISKREILRLNENIKADEDEKMONLSDPFEINVHEDAKEHKEKSVSDRRKRWAG 720
|
|
|
Db 661 AISKREILRLNENIKADEDEKMONLSDPFEINVHEDAKEHKEKSVSDRRKRWAG 720
QY 721 QLVTPDELTLDTSTETREHTVGEYIKLGPNGSPRAMGKSPDVSALKILGEAELOJ 780
|
|
|
Db 721 QLVTPDELTLDTSTETREHTVGEYIKLGPNGSPRAMGKSPDVSALKILGEAELOJ 780
QY 781 ELLENTTIRSEISPEGEKYPRLITGKRVQICISHEINSAIVDSVETKSPSEASPOK 840
|
|
|
Db 781 ELLENTTIRSEISPEGEKYPRLITGKRVQICISHEINSAIVDSVETKSPSEASPOK 840
QY 841 SLKLEGNLEPDDLETEILOEPGSGTNKDESLCTITTDVWISSEKTEKTOGADRIITQEN 900
|
|
|
Db 841 SLKLEGNLEPDDLETEILOEPGSGTNKDESLCTITTDVWISSEKTEKTOGADRIITQEN 900
QY 901 EVSEEGVSTYDOLSDIHIEPGTNDQSHKCDVKSVOPEPFHVKVSEHLNVPOVS 960
|
|
|
Db 901 EVSEEGVSTYDOLSDIHIEPGTNDQSHKCDVKSVOPEPFHVKVSEHLNVPOVS 960
QY 961 VQCSPEESFARSHSLPRKNKNKNSLLGLSTGLFDANPNMLCTCSLPDLSKLFTLM 1020
|
|
|
Db 961 VQCSPEESFARSHSLPRKNKNKNSLLGLSTGLFDANPNMLCTCSLPDLSKLFTLM 1020
QY 1021 DVPYGVRODNLEIDETKDNITKEGPOSEDIYFRETDTLOELQASMOQLRPOGEE 1080
|
|
|
Db 1021 DVPYGVRODNLEIDETKDNITKEGPOSEDIYFRETDTLOELQASMOQLRPOGEE 1080
QY 1081 YSEEEYLVKNSDYPTANGTVADVEDNPNSSSALNEBMSDNGETASCECDSDVEN 1140
|
|
|
Db 1081 YSEEEYLVKNSDYPTANGTVADVEDNPNSSSALNEBMSDNGETASCECDSDVEN 1140
QY 1141 HLEELRLHEOEMGEKFEVEYERIKAIHEDEDNIEICSKYONILGNEHOLYAKTLH 1200
|
|
|
Db 1141 HLEELRLHEOEMGEKFEVEYERIKAIHEDEDNIEICSKYONILGNEHOLYAKTLH 1200
QY 1201 LVMADGAYOEDNDE 1214
|
|
|
Db 1201 LVMADGAYOEDNDE 1214

```

```

; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7144
; SOFTWARE: PL_FL_genes_b Versions 1.0
; SEQ ID NO 2355
; LENGTH: 1243
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-488-725A-2355

Query Match          99.5%; Score 6213; DB 18; Length 1242;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 1213; Conservative 0; Mismatches 1; Indels 28; Gaps 1;

QY 1 MEKVRLQKIGEGSFGKALIVKSTEDGROYVYKEINISRMSSKREESREAVLANMKH 60
|
|
|
Db 1 MEKVRLQKIGEGSFGKALIVKSTEDGROYVYKEINISRMSSKREESREAVLANMKH 60
QY 61 PNIVQYRESEFENGSLYIWDYCEGGDLFRRINAQGVLFQEDDILDMFVOICLAKHVA 120
|
|
|
Db 61 PNIVQYRESEFENGSLYIWDYCEGGDLFRRINAQGVLFQEDDILDMFVOICLAKHVA 120
QY 121 DRLTIHDIKSONFLTKGTVOLGDFGARVINSTVLEARFCIGTPTYYLSPICEKRPY 180
|
|
|
Db 121 DRLTIHDIKSONFLTKGTVOLGDFGARVINSTVLEARFCIGTPTYYLSPICEKRPY 180
QY 181 NKSQDIYALGCYVLELCITKHAPEAGSMKNVLIKIISSFPVSLHYSYDLRSIVSOLF 240
|
|
|
Db 181 NKSQDIYALGCYVLELCITKHAPEAGSMKNVLIKIISSFPVSLHYSYDLRSIVSOLF 240
QY 241 RNPDRPSVNSILEKGTIARIRKIFLSPOLIAEPCITFSFGSOPIPAKRPASGONS 300
|
|
|
Db 241 RNPDRPSVNSILEKGTIARIRKIFLSPOLIAEPCITFSFGSOPIPAKRPASGONS 300
QY 301 SYMPAOKITKPAKYGIPLAYKRYGDKLHEKKPLOKHOAHQTEKRYNNGEERRKISE 360
|
|
|
Db 301 SYMPAOKITKPAKYGIPLAYKRYGDKLHEKKPLOKHOAHQTEKRYNNGEERRKISE 360
QY 361 EAARKRRLEFTEKROKODIISLMKAEOMKROEKERLERINAROGMRNVLASGSGE 420
|
|
|
Db 361 EAARKRRLEFTEKROKODIISLMKAEOMKROEKERLERINAROGMRNVLASGSGE 420
QY 421 VKAPFLGSGGTAPSSFSRGOYEYHAIFDOMOORADNEAKMKREIYGRGLPEROKG 480
|
|
|
Db 421 VKAPFLGSGGTAPSSFSRGOYEYHAIFDOMOORADNEAKMKREIYGRGLPEROKG 480
QY 481 QLAVERAKOVEEFLORRKEAMONKARAGHMYLARLROIRLONFNEROOIKAKLNGEK 540
|
|
|
Db 481 QLAVERAKOVEEFLORRKEAMONKARAGHMYLARLROIRLONFNEROOIKAKLNGEK 540
QY 541 YLAARLROIRLONFNEROOIKAKLNGEKKANHSOGGSEADMRKRIESTLKAHNAARA 600
|
|
|
Db 541 YLAARLROIRLONFNEROOIKAKLNGEKKANHSOGGSEADMRKRIESTLKAHNAARA 600
QY 601 AVLKEDLEKREKREAVEREKRWEEHLVAKGVKSSDVPSPLOHETGSGSPKQOMRSVIV 660
|
|
|
Db 601 AVLKEDLEKREKREAVEREKRWEEHLVAKGVKSSDVPSPLOHETGSGSPKQOMRSVIV 660
QY 633 TSAKKEVGVSSLTDTRETSEMOKTNNAISKREILRLNENIKADEDEKMONLSDPFE 692
|
|
|
Db 633 TSAKKEVGVSSLTDTRETSEMOKTNNAISKREILRLNENIKADEDEKMONLSDPFE 692

```

```

RESULT 3
US-09-488-725A-2355
; Sequence 2355, Application US/09488725A
; GENERAL INFORMATION:
; APPLICANT: Hyseg Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/09/488,725A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21

```

```

Db 661 TSALKEVGDSSLDTRETSSEMOKTNNNAISSKREILRLNENI,KAQOEDKGNLSDF 720
Qy 693 ETINVEDAKHEHEKESVSSDRKKWEAGOLVPLDELITLDTSPSTERTVGEVILGPN 752
Db 721 ETINVEDAKHEHEKESVSSDRKKWEAGOLVPLDELITLDTSPSTERTVGEVILGPN 780
Qy 753 GSPRRAMGKSPDVSILKILGEAELOLOTELLENTTIRSEISPEGGEYKPLITEKKVQCI 812
Db 781 GSPRRAMGKSPDVSILKILGEAELOLOTELLENTTIRSEISPEGGEYKPLITEKKVQCI 840
Qy 813 SHEINPSAIVDSPVETKSPESSEASPOMSKLKGNLEPPDLETELLOEPSGKNKESIP 872
Db 841 SHEINPSAIVDSPVETKSPESSEASPOMSKLKGNLEPPDLETELLOEPSGKNKESIP 900
Qy 873 CITTDVWISSEKETKETSADRTITIOENEVSEGVSTVQDLSDIHIEPTNDOSKCD 932
Db 901 CITTDVWISSEKETKETSADRTITIOENEVSEGVSTVQDLSDIHIEPTNDOSKCD 960
Qy 933 VDKSVQPEPFPHKVVHSEHLNLVPOVOSVOCSPESSEAFRSHSLPPKKNKNSLLIGLS 992
Db 961 VDKSVQPEPFPHKVVHSEHLNLVPOVOSVOCSPESSEAFRSHSLPPKKNKNSLLIGLS 1020
Qy 993 TGLFPANNPKMLRTCSLPLDLSKLFRTLMADVPTVGDVRODMLIDEIKENIKGPDSDSD 1052
Db 1021 TGLFPANNPKMLRTCSLPLDLSKLFRTLMADVPTVGDVRODMLIDEIKENIKGPDSDSD 1080
Qy 1053 IYFEETDIDLOELQASMEQLLRQPGEEYSEESVYKNSDVPPTANGTVDADDEDNPS 1112
Db 1081 IYFEETDIDLOELQASMEQLLRQPGEEYSEESVYKNSDVPPTANGTVDADDEDNPS 1140
Qy 1113 ESALNEEWHNSDNGEISSECEGVNHLLELRLHEQMGKFEFEYERKKAHHEDE 1172
Db 1141 ESALNEEWHNSDNGEISSECEGVNHLLELRLHEQMGKFEFEYERKKAHHEDE 1200
Qy 1173 DENIECSKIYONILGNHOLYAKILHLVMAOGAYOEDNDE 1214
Db 1201 DENIECSKIYONILGNHOLYAKILHLVMAOGAYOEDNDE 1242

```

```

RESULT 4
US-09-783-320-6
; Sequence 6, Application US/09783320
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hu, Yi
; APPLICANT: Nepomichy, Boris
; APPLICANT: Turner, C. Alexander Jr
; APPLICANT: Zambrowicz, Brian
; TITLE OF INVENTION: Novel Human Kinases and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0137-USA
; CURRENT FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: US/09/783,320
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/184,014
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1007
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-783-320-6

```

```

Query Match 82.7%; Score 5162; DB 21; Length 1007;
Best Local Similarity 100.0%; Pred. No. 7,4e-273;
Matches 1007; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 208 MNVLVKIISGSPFVSIHYSDYLSVLSOLFKNRPNDRSPVSIIEKGIAIRIKRPLS 267
Db 1 MNVLVKIISGSPFVSIHYSDYLSVLSOLFKNRPNDRSPVSIIEKGIAIRIKRPLS 60

```

```

Qy 268 POLIAEFCLKTPSPKFGSOPIPAKRPASGONSISVPAOKITTKPAAYGIPLAYKVGDK 327
Db 61 POLIAEFCLKTPSPKFGSOPIPAKRPASGONSISVPAOKITTKPAAYGIPLAYKVGDK 120
Qy 328 KLEHKKPLQKQAOHOTPEKRVNTEGERRKISEEAAARRRLEFIEKKOKQOITISLMA 387
Db 121 KLEHKKPLQKQAOHOTPEKRVNTEGERRKISEEAAARRRLEFIEKKOKQOITISLMA 180
Qy 388 EOKRKOEKLEININRARGWRNVLISAGGSEVAPPLGSGGTIASSPSSROQYEHY 447
Db 181 EOKRKOEKLEININRARGWRNVLISAGGSEVAPPLGSGGTIASSPSSROQYEHY 240
Qy 448 AITDQOQOAEENKAKREITYGSLPRQKQOLAVRAKQVEPLORREANOMARA 507
Db 241 AITDQOQOAEENKAKREITYGSLPRQKQOLAVRAKQVEPLORREANOMARA 300
Qy 508 EGHVYLARLQRILONFNERQOIKAKLGEKKEANHSEGOEGBEADMRRKITESLKAH 567
Db 301 EGHVYLARLQRILONFNERQOIKAKLGEKKEANHSEGOEGBEADMRRKITESLKAH 360
Qy 568 ANRAAVLKEQLERKRRKREAYEREKKWEEHLVAKGVKSSDVSPLQGHETGSPSKQOMR 627
Db 361 ANRAAVLKEQLERKRRKREAYEREKKWEEHLVAKGVKSSDVSPLQGHETGSPSKQOMR 420
Qy 628 SVISVTSALKEVGDSSLDTRETSSEMOKTNNNAISSKREILRLNENI,KAQOEDKGMON 687
Db 421 SVISVTSALKEVGDSSLDTRETSSEMOKTNNNAISSKREILRLNENI,KAQOEDKGMON 480
Qy 688 LSPTFEINVEDAKHEHEKESVSSDRKKWEAGOLVPLDELITLDTSPSTERTVGEVI 747
Db 481 LSPTFEINVEDAKHEHEKESVSSDRKKWEAGOLVPLDELITLDTSPSTERTVGEVI 540
Qy 748 KLEPNSPRAMGKSPDVSILKILGEAELOLOTELLENTTIRSEISPEGGEYKPLITEKK 807
Db 541 KLEPNSPRAMGKSPDVSILKILGEAELOLOTELLENTTIRSEISPEGGEYKPLITEKK 600
Qy 808 KVOCISHEINPSAIVDSPVETKSPESSEASPOMSKLKGNLEPPDLETELLOEPSGKN 867
Db 601 KVOCISHEINPSAIVDSPVETKSPESSEASPOMSKLKGNLEPPDLETELLOEPSGKN 660
Qy 868 DESLPTITDVMISEKETKETOSADRTITIOENEVSEGVSTVQDLSDIHIEPTNDOSQ 927
Db 661 DESLPTITDVMISEKETKETOSADRTITIOENEVSEGVSTVQDLSDIHIEPTNDOSQ 720
Qy 928 HSKCDVDSVQPEPFPHKVVHSEHLNLVPOVOSVOCSPESSEAFRSHSLPPKKNKNSL 987
Db 721 HSKCDVDSVQPEPFPHKVVHSEHLNLVPOVOSVOCSPESSEAFRSHSLPPKKNKNSL 780
Qy 988 LIGLSTGLFPANNPKMLRTCSLPLDLSKLFRTLMADVPTVGDVRODMLIDEIKENIKGP 1047
Db 781 LIGLSTGLFPANNPKMLRTCSLPLDLSKLFRTLMADVPTVGDVRODMLIDEIKENIKGP 840
Qy 1048 SDSEDIYFEETDIDLOELQASMEQLLRQPGEEYSEESVYKNSDVPPTANGTVDADDED 1107
Db 841 SDSEDIYFEETDIDLOELQASMEQLLRQPGEEYSEESVYKNSDVPPTANGTVDADDED 900
Qy 1108 DNPSSEALNEEWHNSDNGEISSECEGVNHLLELRLHEQMGKFEFEYERKKAH 1167
Db 901 DNPSSEALNEEWHNSDNGEISSECEGVNHLLELRLHEQMGKFEFEYERKKAH 960
Qy 1168 IHEDEDENIECSKIYONILGNHOLYAKILHLVMAOGAYOEDNDE 1214
Db 961 IHEDEDENIECSKIYONILGNHOLYAKILHLVMAOGAYOEDNDE 1007

```

```

RESULT 5
US-09-783-320-2
; Sequence 2, Application US/09783320
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hu, Yi
; APPLICANT: Nepomichy, Boris
; APPLICANT: Turner, C. Alexander Jr

```

```
APPLICANT: Zambrowicz, Brian
FILE OF INVENTION: Novel Human Kinases and Polynucleotides Encoding the Same
FILE REFERENCE: LEX-0137-05A
CURRENT APPLICATION NUMBER: US/09/783,320
CURRENT FILING DATE: 2001-02-15
PRIOR APPLICATION NUMBER: US 60/183,582
PRIOR FILING DATE: 2000-02-18
PRIOR APPLICATION NUMBER: US 60/184,014
PRIOR FILING DATE: 2000-02-22
NUMBER OF SEQ ID NOS: 50
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1035
TYPE: PRT
ORGANISM: homo sapiens
US-09-783-320-2
```

```
Query Match      82.3% Score 5138; DB 21; Length 1035;
Best Local Similarity 97.3%; Pred. No. 1,6e-271;
Matches 1007; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

OY 208 MKNVLKTISSGPPVSLHSHYDLSVLSQLEKRNDRPRPSVNSLLEKGFIAKRIEPLS 267
      |||||||
DB 1 MKNVLKTISSGPPVSLHSHYDLSVLSQLEKRNDRPRPSVNSLLEKGFIAKRIEPLS 60

OY 268 POLIAEEFLKTSKSGSOPITPAKRPASGONSISVMPAOKITKPAKYGIPLAYKKYGDK 327
      |||||||
DB 61 POLIAEEFLKTSKSGSOPITPAKRPASGONSISVMPAOKITKPAKYGIPLAYKKYGDK 120

OY 338 KLHEKKPLQKHQAHOPTPEKRVNTGERRKISEEAAKRRRLFEIIEKKOKOQIISLMKA 387
      |||||||
DB 121 KLHEKKPLQKHQAHOPTPEKRVNTGERRKISEEAAKRRRLFEIIEKKOKOQIISLMKA 180

OY 388 EOMKROEKERLERINRAEOGMWNVLSAGSGSEVKAPFLGSGTITAPSSFSRGQYEHYH 447
      |||||||
DB 181 EOMKROEKERLERINRAEOGMWNVLSAGSGSEVKAPFLGSGTITAPSSFSRGQYEHYH 240

OY 448 AIFDMMOQOARADNEAKWKREIYGRGLPEROKGQLAVEAKQVEEFLQKREAMONKARA 507
      |||||||
DB 241 AIFDMMOQOARADNEAKWKREIYGRGLPEROKGQLAVEAKQVEEFLQKREAMONKARA 300

OY 508 EGHM-----YLAIRLQIIRLQNFNEQOIKAKLNGEK 539
      |||||||
DB 301 EGHMIIQNLAMYGGRPSSRGKPRNKEEYVYLARLQIIRLQNFNEQOIKAKLNGEK 360

OY 540 KEANSSEGOESEFADMRKKIESTLKAHANARAAYLKEOLERRKREAYEREKKEWEHLY 599
      |||||||
DB 361 KEANSSEGOESEFADMRKKIESTLKAHANARAAYLKEOLERRKREAYEREKKEWEHLY 420

OY 600 ACYSSSVSPRIGOHETGSGSKOQMSYISTSAKKEGVDSSTLDTRETSEMOKTN 659
      |||||||
DB 421 ACYSSSVSPRIGOHETGSGSKOQMSYISTSAKKEGVDSSTLDTRETSEMOKTN 480

OY 660 MAISSREILRLNENLKAODEKGMONLSDTFEINVEDAKHEKEKSVSDRKKWEG 719
      |||||||
DB 481 MAISSREILRLNENLKAODEKGMONLSDTFEINVEDAKHEKEKSVSDRKKWEG 540

OY 720 GOLVLPDLRLDTSSTTERHNVGEYIKLGPNGSPRRAMGKPTSYVLIIGEAFLQ 779
      |||||||
DB 541 GOLVLPDLRLDTSSTTERHNVGEYIKLGPNGSPRRAMGKPTSYVLIIGEAFLQ 600

OY 780 TELLENTTIRSEISPEGEKRYPLITGEKKVOCISHETNPASIVDSPEVETKSPFESASQ 839
      |||||||
DB 601 TELLENTTIRSEISPEGEKRYPLITGEKKVOCISHETNPASIVDSPEVETKSPFESASQ 660

OY 840 MSILKEGNLEEPDDLETEILOEPSGTNKDESLPCTITDVAIISEKETKETOSADRTTIOE 899
      |||||||
DB 661 MSILKEGNLEEPDDLETEILOEPSGTNKDESLPCTITDVAIISEKETKETOSADRTTIOE 720

OY 900 NEVSEGVSSVTDQSDIHIEGPTNDQSHKCDVAKVQCEPFFHKVHSEHNLVPOVO 959
      |||||||
DB 721 NEVSEGVSSVTDQSDIHIEGPTNDQSHKCDVAKVQCEPFFHKVHSEHNLVPOVO 780
```

```
OY 960 SVQCSPEESFAFRSHSHLPKKNKNKNSLLIGSLGTFDANNPKMLRTCSLPDLSKIFRL 1019
      |||||||
DB 781 SVQCSPEESFAFRSHSHLPKKNKNKNSLLIGSLGTFDANNPKMLRTCSLPDLSKIFRL 840

OY 1020 MDVTVGDVRODULDEIDELKDEINKEGSPDSDDIYEETDIDLOELQASMEQLLREQPGE 1079
      |||||||
DB 841 MDVTVGDVRODULDEIDELKDEINKEGSPDSDDIYEETDIDLOELQASMEQLLREQPGE 900

OY 1080 EYSEESVYAKNSDVEPTANGTDVADEDDNPSSSALNEEMHSDNSDGETIASBECDCSVF 1139
      |||||||
DB 901 EYSEESVYAKNSDVEPTANGTDVADEDDNPSSSALNEEMHSDNSDGETIASBECDCSVF 960

OY 1140 NHLEELRLHLEOEMGEFEVEYEKIKAIHEDEDENIEICKIYONTIGNHOLYAKIL 1199
      |||||||
DB 961 NHLEELRLHLEOEMGEFEVEYEKIKAIHEDEDENIEICKIYONTIGNHOLYAKIL 1020

OY 1200 HLYMADGAYOEDNDE 1214
      |||||||
DB 1021 HLYMADGAYOEDNDE 1035

RESULT 6
US-09-488-725A-5927
; Sequence 5927, Application US/09488725A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/09/488,725A
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7144
; SOFTWARE: PC_FL_genes_b Versions 1.0
; SEQ ID NO 5927
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-488-725A-5927
```

```
Query Match      75.9% Score 4735.5; DB 18; Length 966;
Best Local Similarity 96.3%; Pred. No. 1.3e-249;
Matches 930; Conservative 2; Mismatches 5; Indels 29; Gaps 2;

OY 278 KTFKSGSOPITPAKRPASGONSISVMPAOKITKPAKYGIPLAYKKYGDKILHEKKPLQ 336
      |||||||
DB 1 KTFKSGSOPITPAKRPASGONSISVMPAOKITKPAKYGIPLAYKKYGDKILHEKKPLQ 60

OY 337 KKHQAHOPTPEKRVNTGERRKISEEAAKRRRLFEIIEKKOKOQIISLMKADOMKROEKE 396
      |||||||
DB 61 KKHQAHOPTPEKRVNTGERRKISEEAAKRRRLFEIIEKKOKOQIISLMKADOMKROEKE 120

OY 397 RLEIRINRAEOGMWNVLSAGSGSEVKAPFLGSGTITAPSSFSRGQYEHYHAIPOMOQ 456
      |||||||
DB 121 RLEIRINRAEOGMWNVLSAGSGSEVKAPFLGSGTITAPSSFSRGQYEHYHAIPOMOQ 180

OY 457 RAEDNEAKWKREIYGRGLPEROKGQLAVEAKQVEEFLQKREAMONKARAEGHM----- 511
```

```

Db 181 RAEDNEAKMKREIYGRGLPEROKGOLAVERAKOVEEFLORREKREAOINKRAGCHGILON 240
Qy 512 -----VYLARLQIRLONFNRQOIKAKLRGKEKEANHSQ 548
Db 241 LAAMYGRRSSRGCKPRNKEEYVYLARLQIRLONFNRQOIKAKLRGKEKEANHSQ 300
Qy 549 EGSEADMRKRKTESLKAHANARAVALKEOLERKEAREKREKWEELVAKGVSSDV 608
Db 301 EGSEADMRKRKTESLKAHANARAVALKEOLERKEAREKREKWEELVAKGVSSDV 360
Qy 609 SPPLGQHEHTGSPSKOQMSVISTSAIKVGVDSLTDTRETSEMOKTNNAISSKREI 668
Db 361 SPPLGQHEHTGSPSKOQMSVISTSAIKVGVDSLTDTRETSEMOKTNNAISSKREI 420
Qy 669 LRLNENLKAQDEKQKQNLSDTFEINVEDAKHEKEKSVSSDRKRWAGOLVPLDE 728
Db 421 LRLNENLKAQDEKQKQNLSDTFEINVEDAKHEKEKSVSSDRKRWAGOLVPLDE 480
Qy 729 LTLDTSESTERHTVGEVILKIPNGSPRRAMGKSPDVSYLKILGEAELOLOTELLENTTI 788
Db 481 LTLDTSESTERHTVGEVILKIPNGSPRRAMGKSPDVSYLKILGEAELOLOTELLENTTI 540
Qy 789 RSEISPEGEKKPLITGEKKVOCISHEINPSAIVDSPVETKSPFESEASPOMSLKLEGNL 848
Db 541 RSEISPEGEKKPLITGEKKVOCISHEINPSAIVDSPVETKSPFESEASPOMSLKLEGNL 600
Qy 849 EEPDDELETLQEPSTGNKDESLPCTITDVMISEKETKETOSADRITTOENEVEDGVS 908
Db 601 EEPDDELETLQEPSTGNKDESLPCTITDVMISEKETKETOSADRITTOENEVEDGVS 660
Qy 909 STVQOLSDIHEPTGNDOSHKCDVKSVPPEPFKVVHSEHLNLPVOVQVOCSPSEES 968
Db 661 STVQOLSDIHEPTGNDOSHKCDVKSVPPEPFKVVHSEHLNLPVOVQVOCSPSEES 720
Qy 969 FAFRSHSHLPKKNKNSLLIGSLGTFDANNKMLRTCSLPDLSKLFRTLMQVPTVGDV 1028
Db 721 FAFRSHSHLPKKNKNSLLIGSLGTFDANNKMLRTCSLPDLSKLFRTLMQVPTVGDV 780
Qy 1029 RODMLEIDELIKENIKGSPDSSEDIYFEETDLOELQASMDQLKEOGEEYSEESBV 1088
Db 781 RODMLEIDELIKENIKGSPDSSEDIYFEETDLOELQASMDQLKEOGEEYSEESBV 840
Qy 1089 LKNSDVEPTANGTVDADEDNPSSESALNEMHSDNSDEIASSECECDSEVFNHELERLH 1148
Db 841 LKNSDVEPTANGTVDADEDNPSSESALNEMHSDNSDEIASSECECDSEVFNHELERLH 900
Qy 1149 LEOEMGEKFEVEYERIKAIHEDEDENIETISKIVONITIGNEHOHLYAKILHLVADGAY 1208
Db 901 LEOEMGEKFEVEYERIKAIHEDEDENIETISKIVONITIGNEHOHLYAKILHLVADGAY 960
Qy 1209 QEDNDE 1214
Db 961 QEDNDE 966

```

```

; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7144
; SOFTWARE: PL_FL_genes_b Versions 1.0
; SEQ ID NO: 5928
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-488-725A-5928

Query Match 75.9%; Score 4735.5; DB 18; Length 966;
Best Local Similarity 96.3%; Pred. No. 1,36-249;
Matches 930; Conservative 2; Mismatches 5; Indels 29; Gaps 2:

Qy 278 KTFPSKFGSOPi-PAKRPAQONSISVMPAOKITTKPAKYGIPLAYKKYGDKKLHEKKPLQ 336
Db 1 KHFSKFGSALYOLKRPASGONSISVMPAOKITTKPAKYGIPLAYKKYGDKKLHEKKPLQ 60
Qy 337 KHKQAHQTEPEKRYNVTGEERKISEEAKRRRLFEFEKEKKOKDOISLMAEQMKROEKE 396
Db 61 KHKQAHQTEPEKRYNVTGEERKISEEAKRRRLFEFEKEKKOKDOISLMAEQMKROEKE 120
Qy 397 RLERIRAREOGRRNVLASGSGEYKAPFLGSGGTIAPSSFSRGQYEHYAIPOMOQ 456
Db 121 RLERIRAREOGRRNVLASGSGEYKAPFLGSGGTIAPSSFSRGQYEHYAIPOMOQ 180
Qy 457 RAEDNEAKMKREIYGRGLPEROKGOLAVERAKOVEEFLORREKREAOINKRAGCHGILON 511
Db 181 RAEDNEAKMKREIYGRGLPEROKGOLAVERAKOVEEFLORREKREAOINKRAGCHGILON 240
Qy 512 -----VYLARLQIRLONFNRQOIKAKLRGKEKEANHSQ 548
Db 241 LAAMYGRRSSRGCKPRNKEEYVYLARLQIRLONFNRQOIKAKLRGKEKEANHSQ 300
Qy 549 EGSEADMRKRKTESLKAHANARAVALKEOLERKEAREKREKWEELVAKGVSSDV 608
Db 301 EGSEADMRKRKTESLKAHANARAVALKEOLERKEAREKREKWEELVAKGVSSDV 360
Qy 609 SPPLGQHEHTGSPSKOQMSVISTSAIKVGVDSLTDTRETSEMOKTNNAISSKREI 668
Db 361 SPPLGQHEHTGSPSKOQMSVISTSAIKVGVDSLTDTRETSEMOKTNNAISSKREI 420
Qy 669 LRLNENLKAQDEKQKQNLSDTFEINVEDAKHEKEKSVSSDRKRWAGOLVPLDE 728
Db 421 LRLNENLKAQDEKQKQNLSDTFEINVEDAKHEKEKSVSSDRKRWAGOLVPLDE 480
Qy 729 LTLDTSESTERHTVGEVILKIPNGSPRRAMGKSPDVSYLKILGEAELOLOTELLENTTI 788
Db 481 LTLDTSESTERHTVGEVILKIPNGSPRRAMGKSPDVSYLKILGEAELOLOTELLENTTI 540
Qy 789 RSEISPEGEKKPLITGEKKVOCISHEINPSAIVDSPVETKSPFESEASPOMSLKLEGNL 848
Db 541 RSEISPEGEKKPLITGEKKVOCISHEINPSAIVDSPVETKSPFESEASPOMSLKLEGNL 600
Qy 849 EEPDDELETLQEPSTGNKDESLPCTITDVMISEKETKETOSADRITTOENEVEDGVS 908
Db 601 EEPDDELETLQEPSTGNKDESLPCTITDVMISEKETKETOSADRITTOENEVEDGVS 660
Qy 909 STVQOLSDIHEPTGNDOSHKCDVKSVPPEPFKVVHSEHLNLPVOVQVOCSPSEES 968
Db 661 STVQOLSDIHEPTGNDOSHKCDVKSVPPEPFKVVHSEHLNLPVOVQVOCSPSEES 720
Qy 969 FAFRSHSHLPKKNKNSLLIGSLGTFDANNKMLRTCSLPDLSKLFRTLMQVPTVGDV 1028
Db 721 FAFRSHSHLPKKNKNSLLIGSLGTFDANNKMLRTCSLPDLSKLFRTLMQVPTVGDV 780

```

```

RESULT 7
US-09-488-725A-5928
; Sequence 5928, Application US/09488725A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/09/488,725A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19

```



```

OY      421 VKFFLGSGTIPISFSSRSGOYEYHAIIPOMOOOAEDEAKMKREIYGRGLPEROKG      480
Db      420 VKASFEGIGGAVSPSPCSRGROYEHYHAIIPDMQRLRAEDNEAMKCGIYGRLLPEROKG      479
OY      481 OLAVERRKOYEETLQKREAMONKRAEGHWYLLARLQRLQNFQNFNRQOIKAKLGEKK      540
Db      480 HLAVERANQOYEETLQKREAMONKRAEGHWYLLARLQRLQNFQNFNRQOIKAKLGEKK      539
OY      541 EAHNSEOESSEADMRKKKISTSLAHANARAYLKLQLEKKREAEYERKKWEEHLYA      600
Db      540 EADDTKQGEYETLTDHKLKMESTLAAQTNARAYLKLQLEKKREAEYERKKWEEHLYA      599
OY      601 KGVKSSDYSPPLCGHETG      618
Db      600 R-VKSSDVPLELLELTG      616

RESULT      10
US-09-873-404-4
? Sequence 4, Application US/09873404
? GENERAL INFORMATION:
? APPLICANT: WEBSTER, Marion et al
? TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
? TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
? TITLE OF INVENTION: THEREOF
? FILE REFERENCE: CL001212-CIP
? CURRENT APPLICATION NUMBER: US/09/873,404
? NUMBER OF SEQ ID NOS: 4
? SOFTWARE: FastSeq for Windows Version 4.0.
? SEQ ID NO 4
? LENGTH: 616
? TYPE: PRT
? ORGANISM: Mouse
? US-09-873-404-4

```

QY	481	QVVEAKQVVEEFLORKEAMONAKRABGVHVALRDRITLONNEQOIKALREKK	54.0
Db	480	HLAVEAQAQVEEFLORKEAMONAKRABGVHVALRDRITLONNEQOIKALBREKK	53.9
QY	541	EAAHSGGSGSSEAAEMRRKKTLESLKAAHNAARVALKEOLERRKRAAREKKVMEHLVA	600
Db	540	EAAHTGGGQATEETMTLRKKMESLKAQTNAAVALKEOLERRKRAAREKKVMEHLVA	599
QY	601	KGVKSSDVSPGLQGHETG	618
Db	600	R-VKSSDVPLPLELLETG	616

```

RESULT 11
US-09-870-962-6
Sequence 6: Application US/09870962
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Tang, Y. Tom
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
APPLICANT: Gorgone, Gina
APPLICANT: Azimzal, Valda
APPLICANT: Lu, Alina
TITLE OF INVENTION: Protein Kinase Homologs
FILE REFERENCE: PF-0614 US
CURRENT APPLICATION NUMBER: US/09/870,962
CURRENT FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: 09/420,915
PRIOR FILING DATE: 1999-10-20
PRIOR APPLICATION NUMBER: US 09/173,581
PRIOR FILING DATE: 1998-10-15
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PERL Program
SEQ ID NO 6
LENGTH: 345
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE: -
US-09-870-962-6
OTHER INFORMATION: 1567782

```

	Query Match	Similarity	44.68%	Score 2785	DB 22	Length 616
	Best Local	Similarity	88.08%	Pred. No 2, 2c-143		
	Matches	544;	Conservative	27;	Mismatches	45;
					Indels	2;
					Gaps	2;
QY	1	METVYRLQRTGEGSEPEKALITVSTEDGYOYAIKEINISRMSSKREBSREVVAVIANMMH	60			
Db	1	METVYRLQRTGEGSEPEKALVAVSTEDGHHYAIKEINISRMSSKREBSREVVAVIANMMH	60			
QY	61	PNIVQYRESPEENGSLYIVMDYCEGGDLPFKYINAKOYVLEOEOIILDMFOICLAKHYH	120			
Db	61	PNIVQYKESPEENGSLYIVMDYCEGGDLPFKYINAKOYVLEOEOIILDMFOICLAKHYH	120			
QY	121	DKRIILHDKTQSNMIFLTKDCGYQLDDECLAVIANSYELATCIGTGYIYLSPEICENKRY	180			
Db	121	DKRIILHDKTQSNMIFLTKDCGYQLDDECLAVIANSYELATCIGTGYIYLSPEICENKRY	180			
QY	181	NNKSDIATGACGYIYEICTLKAHAFEGNMKNVYKLTISGPPSPSLHNSYLSRSLVQLRK	240			
Db	181	NNKSDIATGACGYIYEICTLKAHAFEGNMKNVYKLTISGPPSPSPHSDILSRSLVQLRK	240			
QY	241	RNPDRDPVSNSITLKEFKIAKRIEKLSQLIAEPECLKTSKGSQCPITPAKKPAAQSONSI	300			
Db	241	RNPDRDPVSNSITLKEFKIAKRIEKLSQLIAEPECLKTSKGSQCPITPAKKPAAQSOVS	300			
QY	301	SVNPAQKTIIPKAAKTYIPLATYKKTGDKKLHEKKPLQKHQAHOQTPEKRVNTGEERRKISE	360			
Db	301	SEVPAAQKTIIPKAAKTYVPLATYKKTGDKKLHEKKPLPKKHQAHOQPVAKNMSSGERKKRME	360			
QY	361	EAARKRRLEFIEKKKQKQDIIISLKAADQMRQEKERHERINRAAREGMMNVLSAGSGSE	420			
Db	361	EAARKRRLEFIEKKKQKQDQ - IRLFKAAQMRQEKERHERINRAAREGMMNVLRAGSGSE	419			
QY	421	VKAPFLGSGGTIAPSSFSRGOYEYHYHAIIPDOMQOQRAEDNAEKMKREIYGRGLPRQGG	480			
Db	420	VKASFEIGGAVSPSCSPRGQYEHYHAIIPDOMQRLRAEDNAEWKQSIGRMLPRQGG	479			

Query Match	27.88	Score 1734	DB 22	Length 345
Best Local Similarity	90.55	Pred. No. 2	4e-86	
Matches 341	Conservative 3	Mismatches 1	Indels 32	Gaps
Qy	1	METVYALQKIGTGGSGCKILVKSFTEDBROVYVTEINISMSSSKEKEEESRRVAVYAMMKH	60	
Db	1	METVYALQKIGTGGSGCKILVKSFTEDBROVYVTEINISMSSSKEKEEESRRVAVYAMMKH	60	
Qy	61	PNVVOYRESEENSGLYVMDYCGGGLFKNIAQGVLFQEDOLLDFVQICLALNVH	120	
Db	61	PNVVOYRESEF-----GLIDMFVQICLALNVH	88	
Qy	121	DKRIILHRDKSONIFLTEDGTVOIGDEGIAFVNVSTVELAPFCITPPVYSPEICEKPY	180	
Db	89	DKRIILHRDKSONIFLTEDGTVOIGDGIKAVNVSTVELAPFCITPPVYSPEICEKPY	148	
Qy	181	NNKSDIWAIGCVLYELCTLTKIAFAFGSMKMLVLTITSSGFPVYSLSHYSTDURSLVSOLF	240	
Db	149	NNKSDIWAIGCVLYELCTLTKIAFAFGSMKMLVLTITSSGFPVYSLSHYSTDURSLVSOLF	208	
Qy	241	RNRDRPVSNIILEKGFIAKKRIEKFISPOLIAAEFLCTFKFSFGSOPIPARRPASGNSI	300	
Db	209	RNRDRPVSNIILEKGFIAKKRIEKFISPOLIAAEFLCTFKFSFGSOPIPARRPASGNSI	268	
Qy	301	SVMPAQKIRTPAAKAGIGILAANKKYGDKMLKHEKPPLOKNAQQAQTEKRVNTGEERRKISE	360	
Db	269	SVMPAQKIRTPAAKAGIGILAANKKYGDKMLKHEKPPLOKNAQQAQTEKRVNTGEERRKISE	328	


```
QY      361 EAARKRLEFIEEKKQ 377
        |||||
Db      329 EAARKRLEFIEKDKER 345
```

RESULT 12
 US-09-715-427-35
 Sequence 35, Application US/09715427
 GENERAL INFORMATION:
 APPLICANT: Quinn, Kerry E.
 APPLICANT: Speitek, Kimberly A.
 APPLICANT: Majumder, Kumud
 APPLICANT: Vermet, Corine
 APPLICANT: Herrmann, John L.
 APPLICANT: Burgess, Catherine
 APPLICANT: Fernandes, Elma
 APPLICANT: Taupier Jr., Raymond
 APPLICANT: Rastelli, Luca
 APPLICANT: Curagen Corporation
 TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
 FILE REFERENCE: 15966-598
 CURRENT FILING DATE: 2000-11-16
 PRIOR APPLICATION NUMBER: US/09/715,427
 PRIOR FILING DATE: 1998-11-13
 PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,839
 PRIOR FILING DATE: 2000-04-05
 PRIOR APPLICATION NUMBER: U.S.S.N. 60/195,637
 PRIOR FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: U.S.S.N. 60/197,080
 PRIOR FILING DATE: 2000-04-13
 PRIOR APPLICATION NUMBER: U.S.S.N. 60/232,677
 PRIOR FILING DATE: 2000-09-15
 PRIOR APPLICATION NUMBER: U.S.S.N. 60/181,347
 PRIOR FILING DATE: 2000-02-09
 PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,195
 PRIOR FILING DATE: 2000-04-03
 NUMBER OF SEQ ID NOS: 53
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO 35
 LENGTH: 254
 TYPE: PRT
 ORGANISM: Mus musculus
 US-09-715-427-35

	Query Match	Score	1283:	DB 21:	Length	254:			
	Best Local Similarity	96.5%:	Pred. No. 6,7e-62:						
	Matches 245:	Conservative	5:	Mismatches	4:	Indels	0:	Gaps	0:
OY	5	VRLQKIGESFGKALLIVKSTEDGROYVKEINISMSKSEPEESRRRVAYLANMKKHNTY	64						
Db	1	VRLQKIGESFGKAVLYKSTEDGRHVIKEINISMSKKEQESREVVAYLANMKKHNTY	60						
OY	65	OYRSEPEENGSLTYWDYCEGGDLFRKRIINAQGVYFQEDQILLMFVQICAIKHYHDRKI	124						
Db	61	OYKSEPEENGSLTYWDYCEGGDLFRRIINAQGVYFQEDQILLMFVQICAIKHYHDRKI	120						
OY	125	LHRDIKQNTFLTKDGTVOLGDBGFARVYNSVTEIARFCITGPYYISLPCEIKENPKYNNKS	184						
Db	121	LHRIRKQNTFLTKDGTVOLGDBGFARVYNSVTEIARFCITGPYYISLPCEIKENPKYNNKS	180						
OY	185	DYALGCVLYELCTLKAHAFBAGSMKNLVYIKITSSGFPVSLHYSTDRLSYQLFKENPR	244						
Db	181	DYALGCVLYELCTLKAHAFBAGSMKNLVYIKITSSGFPVSPHYSTDRLSYQLFKENPR	240						
OY	245	DRPSVNSILKEGFI	258						
Db	241	DRPSVNSILKEGFI	254						

RESULT 13
US-09-898-837A-35

```

1 Sequence 35 Application us/09898837A
2 GENERAL INFORMATION:
3 APPLICANT: Quinn, Kerry E.
4 APPLICANT: Spytek, Kimberly A.
5 APPLICANT: Majumder, Kumud
6 APPLICANT: Vernet, Corine
7 APPLICANT: Herrmann, John L.
8 APPLICANT: Burgess, Catherine
9 APPLICANT: Fernandes, Elma
10 APPLICANT: Taupier Jr., Raymond
11 APPLICANT: Rastelli, Luca
12 APPLICANT: Cuiagen Corporation
13 APPLICANT: Gerlach, Valerie L
14 APPLICANT: MacDougall, John R
15 TITLE OF INVENTION: NOVEL SERINE/THREONINE PROTEIN-KINASE LIKE PROTEINS AND
16 FILE OF INVENTION: NUCLEIC ACIDS ENCODING THE SAME
17 FILE REFERENCE: 15966-598 CIP
18 CURRENT APPLICATION NUMBER: us/09/898, 837A
19 PRIOR FILING DATE: 2001-07-03
20 PRIOR APPLICATION NUMBER: U.S.S.N. 60/165, 986
21 PRIOR FILING DATE: 1999-11-17
22 PRIOR APPLICATION NUMBER: U.S.S.N. 60/194, 839
23 PRIOR FILING DATE: 2000-04-05
24 PRIOR APPLICATION NUMBER: U.S.S.N. 60/195, 637
25 PRIOR FILING DATE: 2000-04-07
26 PRIOR APPLICATION NUMBER: U.S.S.N. 60/197, 080
27 PRIOR FILING DATE: 2000-04-13
28 PRIOR APPLICATION NUMBER: U.S.S.N. 60/232, 677
29 PRIOR FILING DATE: 2000-09-15
30 PRIOR APPLICATION NUMBER: U.S.S.N. 60/181, 347
31 PRIOR FILING DATE: 2000-02-09
32 PRIOR APPLICATION NUMBER: U.S.S.N. 60/194, 195
33 PRIOR FILING DATE: 2000-04-03
34 PRIOR APPLICATION NUMBER: U.S.S.N. 60/215, 906
35 PRIOR FILING DATE: 2000-07-03
36 PRIOR APPLICATION NUMBER: U.S.S.N. 09/715, 427
37 PRIOR FILING DATE: 2000-11-16
38 NUMBER OF SEQ ID NOS: 53
39 SOFTWARE: Patencin Ver. 2.1
40 SEQ ID NO 35
41 LENGTH: 254
42 TYPE: PRT
43 ORGANISM: Mus musculus
44 US-09-898-837A-35

```

Query Match	20.6%	Score 1283	DB 22	Length 254
Best Local Similarity	96.5%	Pred. No. 6.7e-62		
Matches 245	Conservative	5	Mismatches 4	Indels 0
QY	5	VRLQITGGSGGATITVSTEDGROVYIKREITISMSKSEKEEESRREAYAVLANKHPNIV	64	
DB	1	VRLQITGGSGGAAVLKSTEDGRHYIKEITISMSKSEKEEESRREAYAVLANKHPNIV	60	
QY	65	QYRSPEENSIXIYWDYCGGDLFRTIAQGVLFQEDQIILDMVQICLAKHVHDKI	124	
DB	61	QYRSPEENSIXIYWDYCGGDLFRTIAQGVLFQEDQIILDMVQICLAKHVHDKI	120	
QY	125	LHRPIKSONIFLRKDDGYVQDGRGTAIVANSVTEIARICIGPPYLSPEICENKPYNNKS	184	
DB	121	LHRPIKSONIFLRKDDGYVQDGRGTAIVANSVTEIARICIGPPYLSPEICENKPYNNKS	180	
QY	185	DIALAGVLYELCTIKHAFAGSMKMLVYKIIISGSPPYSLHYSVDLSLVSOLFKNRPR	244	
DB	181	DIALAGVLYELCTIKHAFAGSMKMLVYKIIISGSPPYSLHYSVDLSLVSOLFKNRPR	240	
QY	245	DRPSVNSILKEGPT	258	
DB	241	DRPSVNSILKEGPT	254	

RESULT 14
US-09-760-446A-1590

Wed May 15 14:13:57 2002

us-09-783-320-4.tapm

Page 10

```
Sequence 1590, Application US/09760446A
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PTZ68
CURRENT APPLICATION NUMBER: US/09/760,446A
CURRENT FILING DATE: 2000-01-16
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/225,757
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/226,868
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/216,647
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,267
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/216,880
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,270
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/251,869
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/235,834
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/234,274
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/234,223
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/228,924
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/224,518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224,519
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241,809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249,299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236,327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241,785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08

PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234,997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237,039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/236,370
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/236,802
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,037
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240,960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239,935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239,937
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/241,787
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,474
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,532
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,216
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,210
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/226,681
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,759
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225,213
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/227,182
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,214
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225,266
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249,218
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,208
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,213
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,212
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,207
```

PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,245
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,244
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,217
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,211
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,215
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,264
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,214
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,297
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/232,400
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/231,242
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,081
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,080
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,414
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,244
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,064
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/233,063
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,397
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,399
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,401
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,808
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,826
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,786
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,221
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,475
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/231,243
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,065
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,398

Query Match 18.4% Score 1146; DB 21; Length 256;
Best Local Similarity 99.1% Pred. No. 2e-54; Mismatches 0; Indels 0; Gaps 0;
Matches 220; Conservative 2;

QY 153 LNSTELARTCIGTPYLLSPETICENKPNKSDIMALGCVLYELCTLKAHAFAGSMKNLY 212
DB 35 LNSTELARTCIGTPYLLSPETICENKPNKSDIMALGCVLYELCTLKAHAFAGSMKNLY 94
QY 213 LKTIISGPPVSLHYSYDLRSVLSQLFKRNPDRPVSNSILEKGFIAKRIEKFLEPQLA 272
DB 95 LKTIISGPPVSLHYSYDLRSVLSQLFKRNPDRPVSNSILEKGFIAKRIEKFLEPQLA 154
QY 273 EEECLKTFKFSQSOPIPAKRPASGONSISVPAOKIRPKPAKYGIPLAVAKKYGDKKLHEK 332
DB 155 EEECLKTFKFSQSOPIPAKRPASGONSISVPAOKIRPKPAKYGIPLAVAKKYGDKKLHEK 214
QY 333 KPIOKKHOHPPEKRVNTGEERRKISEEAAKRRLEFIKE 374

DB 215 KPIOKKHOHPPEKRVNTGEERRKISEEAAKRRLEFIKEK 256

RESULT 15
US-09-824-583-2
Sequence 2, Application us/09824583
GENERAL INFORMATION:
APPLICANT: WEBSTER, Marion et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: C1001212
CURRENT APPLICATION NUMBER: US/09/824,583
CURRENT FILING DATE: 2001-04-03
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 2
LENGTH: 640
TYPE: PRT
ORGANISM: Human
US-09-824-583-2

Query Match 17.0% Score 1060; DB 22; Length 640;
Best Local Similarity 33.0% Pred. No. 3.5e-49;
Matches 270; Conservative 128; Mismatches 203; Indels 218; Gaps 22;

QY 1 MEKVYRLQIOEGSGKAIIVKSTEDGROYVKEINISMSKERESESRREAVIANMKH 60
DB 1 MDKIDVIAKIQGAFKAYLAKGSKDVKIENIFEMPIOEKASKEKRVILLKMKH 60
QY 61 PRVIOYRSPFENGSLYIVMDYEGGDLFKRIAQGVLFQDDIILDFVOICLALKHVH 120
DB 61 PRVIAFNSFQENGRRLTIVMEYCDGDLKMRINRQGVLFSEDDILGVEVOISIGLKH 120
QY 121 DRKILHDIKSONFLTKDGTV-OLDFGIARVNSTVLAFTGTPYLLSPETICENK 179
DB 121 DRKILHDIKSONFLTKDGTV-OLDFGIARVNSTVLAFTGTPYLLSPETICENK 180
QY 180 YNKSIDWALCCVLYELCTLKAHAFAGSMKNLYLKTISSPFPVSLHYSYDLRSYQLF 239
DB 181 YNKSIDWALCCVLYELCTLKAHAFAGSMKNLYLKTISSPFPVSLHYSYDLRSYQLF 240
QY 240 KNPDRPVSNSILEKGFIAKRIEKFLEPQLAEEFCLTFSKFSQOPIPAKRPASGONS 299
DB 241 QVSPDRPVSNSILEKGFIAKRIEKFLEPQLAEEFCLTFSKFSQOPIPAKRPASGONS 293
QY 300 ISVPAOKIRPKPAKYGIPLAVAKKYGDKKLHEKRPLOKHOHPPEKRVNT 351
DB 300 ISVPAOKIRPKPAKYGIPLAVAKKYGDKKLHEKRPLOKHOHPPEKRVNT 351
QY 294 -KVVOCKKIQVREFGKCPFRSISVPI-----KRAVILHREWRP----- 334
DB 352 GEERRKISEEAAKRRLEFIKEKKOKDQIISLMAKQMKROEKRLERINRAREQWRN 411
DB 335 -----AGQKARSIMIER----- 348
QY 412 VLSAGSGEVKAPFLGSGGTIAPSSFSRCOYEHYNAIFDOMOQRAEDN-----EAKWK 466
DB 349 -----PRLAIVCGHYIDYYNOJLDMRLRRHKRSYHPIQENTG 386
QY 467 RELYNGGLPEBROGOLAVERAQOYEFLORREMOAKAEBHAYYLARLQIRQO--- 523
DB 387 VEDYQ-----QETHGHSPPQWPAEYLQKRFPAQYKLVK-----KOLGRSS 431
QY 524 -----NENRQITAKILRGKKEKKAHNSGEGSEEDMMKRIEELKANAARAVLKEQLE 580
DB 432 AEBNVMQROELR-----SNGFE-----PFOQLPRKN-----E 460
QY 581 RKRREAYEREKKWEEHLVAVGKSSDVSPPLGQHEFGSPSKQOORRSYISYALKEVG 640
DB 461 MKQOEYWKQLEIRQOYL-----NDMKETIRKMGEPEDIDKLOKMK----- 503
QY 641 VDSLTDIETSEEMOKTNNALISKRTELRLN-----ENLKAODEKMONLSDTIE 693

Wed May 15 14:13:57 2002

us-09-783-320-4.ram

Page 12

Db 504 ----LQNTKESKNPEOK---YKAKGVKFEINLDKISDENILOEEBAMIPNETLFE 555
QY 694 ----INVHEDAKHEHEKESVSDDRKKW-EAGGOLVIPLELTLDTSF-----STERH 741
Db 556 DGMKFEYECVKEHGDYTDKAFELHCPFAAFTELTLWSFLFLEXSLPHFILEKSPFSRH 615
QY 742 TVGEVIRLGPNGSPRAWGKSPDVSUKIILGEAEILOOT 780
Db 616 LIEDLL-----CANDCSLKDMSKEMELRT 640

Search completed: May 15, 2002, 07:59:35
Job time: 395 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 15, 2002, 06:30:42 ; Search time 128.31 Seconds
(without alignments)

1050.920 Million cell updates/sec

Title: US-09-783-320-4

Perfect score: 6243
Sequence: 1 MEKTVRLQKIGESFGKAIL.....YAKIHLVMADGAYQEDNDE 1214

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /net/abs06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6243	100.0	1214	AAU07102	Human novel human
2	6237	99.9	1214	AAAM39211	Human polypeptide
3	6213	99.5	1242	AAAM39210	Human polypeptide
4	5162	82.7	1007	AAU07103	Human novel human
5	5138	82.3	1035	AAU07101	Human novel human
6	4735.5	75.9	966	AAAM40996	Human polypeptide
7	4735.5	75.9	966	AAAM40997	Human polypeptide
8	1734	27.8	345	AAAT6753	Human protein kina
9	1734	27.8	345	AAE06211	Human protein kina
10	1387	22.2	320	AAAB9221	Human protein sequ
11	1025	16.4	649	AAU03545	Human protein kina

12	882	14.1	506	22	AAAT78344	Human protein SEQ
13	875	14.0	527	22	AAAT79328	Human protein SEQ
14	806.5	12.9	510	21	AAAB68778	Amino acid sequenc
15	628	10.1	658	22	AAAB65702	Novel protein kina
16	604	9.7	645	22	AAU03546	Human protein kina
17	599.5	9.6	979	19	AAAM40309	Human ITAK protein
18	599.5	9.6	979	21	AAAY9294	Human GEF containi
19	593.5	9.5	836	22	AAAB65703	Novel protein kina
20	584	9.4	841	22	AAAB65338	Drosophila melanog
21	557	8.9	330	22	AAAB85038	Human SERA protein
22	550	8.8	330	22	AAAT92330	Human NIK1 protein
23	538	8.6	336	21	AAAT92148	Human NERK protein
24	519	8.3	360	21	AAAT92147	Drosophila melanog
25	503	8.1	360	21	AAAT92147	E. nidulans NIMA p
26	502.5	8.0	1142	22	AAAG70713	S. cerevisiae apopt
27	494.5	7.9	1298	22	AAAB68222	Amino acid sequenc
28	494	7.9	1353	22	AAAB68219	Amino acid sequenc
29	493.5	7.9	1202	21	AAAB08521	A murine Ste20-rel
30	492	7.9	1339	22	AAAT78421	Human protein SEQ
31	492	7.9	1385	22	AAAT79405	Human protein SEQ
32	490.5	7.9	1306	22	AAAB68218	Amino acid sequenc
33	490	7.8	1269	22	AAAB68223	Human protein kina
34	490	7.8	1360	21	AAAB50059	Large NIK-Related
35	487.5	7.8	1324	22	AAAB68221	Amino acid sequenc
36	486	7.8	1277	22	AAAB68220	Amino acid sequenc
37	485.5	7.8	302	21	AAAB68220	Human cardiovascu
38	485.5	7.8	302	21	AAAT59143	Human TGF-beta rec
39	485.5	7.8	302	22	AAAT59143	Human TGF-beta rec
40	485.5	7.8	311	22	AAAT59143	Human TGF-beta rec
41	485.5	7.7	311	22	AAAT59143	Human TGF-beta rec
42	485.5	7.7	311	22	AAAT59143	Human TGF-beta rec
43	478.5	7.7	302	22	AAAB68217	Amino acid sequenc
44	478.5	7.7	302	22	AAAT59146	Human serine/threo
45	477.5	7.6	313	21	AAAB50054	Murine Dendritic C
					AAAB22799	Human nek4 protein

ALIGNMENTS

RESULT	1
AAU07102	AAU07102 standard; Protein: 1214 AA.
ID	AAU07102
XX	AAU07102;
AC	AAU07102;
XX	24-OCT-2001 (first entry)
DT	XX
XX	Human novel human protein, NHP #2.
DE	XX
XX	Human: novel human protein; NHP; breast cancer; prostate cancer;
KW	Immunogen; antibody; gene therapy; antisense.
XX	XX
OS	Homo sapiens.
XX	XX
PN	WO200161016-A2.
PD	23-AUG-2001.
XX	XX
PF	15-FEB-2001; 2001WO-US05356.
XX	XX
PR	18-FEB-2000; 2000US-0183582.
PR	22-FEB-2000; 2000US-0184014.
XX	XX
PA	(LEXI-) LEXICON GENETICS INC.
XX	XX
FI	Walke DW, Hu Y, Nepomichy B, Turner CA, Zambrowicz B;
XX	XX
XX	WPI; 2001-602793/55.
DR	N-PSDB; AAS11558.
XX	XX
PT	Isolated nucleic acids encoding novel human proteins useful for the
XX	treatment of disease and as probes for testing and detection -

PS Claim 4; Page 37-39; 69pp; English.

XX The invention relates to novel human proteins (NHP) and the nucleic acids encoding them. The nucleic acids encode mammalian transporter proteins and are useful for the treatment (e.g. by gene therapy or antisense technology) of any of a wide variety of symptoms associated with biological disorders (e.g. breast and prostate cancer) or imbalances and as probes for the identification, selection and validation of novel molecular targets for drug discovery. The proteins may be used to raise anti-NHP antibodies. The present sequence represents an NHP of the invention.

XX Sequence 1214 AA:

Query Match 100.0%; Score 6243; DB 22; Length 1214;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKYYRLQKIGSGFGKAILVKSTEDGRQYVKEINISRMSSKERESEREAVLANMKH 60
DB 1 mekyvrlqkigsgfghkailvkstedyqvyikeinismsskereesrrevavlammkh 60
QY PINIVYRESFEENGSLYIVMDYCEGGDLFRINAKGVLFQEDQILDWFQICLALKHYH 120
DB 61 pinivvresfeengslyivmdyceggdlfrinaqgvllfqedqildwfvqiclaalkhvh 120
QY 121 DKRIIHRDIKSONIFLTGDTVOLGDFGIARVLNSTVELARTCGIPPYLSPETICENKPY 180
DB 121 dkrilhrdiksonifltgdtvolgdfgiarvlnstvelartcgitppylspeticeknp 180
QY 181 NKRSDIMALGVLYELCTLKAHAFEGSMKNLYIKTSGSPPSYLAHSTYLSRLVQLTK 240
DB 181 nksrdimalgvlyelctlkahafegsmknlyiksgsppsylahstysrlvqltk 240
QY 241 RNPDRPSVNSTLEKGFIAKRIEFLSPQILAEFCLKTSGSGOPITPAKRPAQONS 300
DB 241 rnpdrpsvnsstlekgfiaakrie flspqilaeefclktsgsgopitpakra paqons 300
QY 301 SVMPAQKITPAKAYGIDPLAVKRYGDKKLHEKKPLQKHKAHQHOPTEKRVNTGERRKIS 360
DB 301 svmpaqkitlpaakayidplavkkygdkklhekplqkhkqhahqptekrvntgeerrkise 360
QY 361 EAARRRRLFEITEKERRQDQITSLMKAQOMKROEKERLERINRAEQGMHNVLSAGSGS 420
DB 361 eaarrrrlfeitekerrkdqit slmkaeqomkroekerlerinraeqgmh nvlvsagsgs 420
QY 421 VKAPFLGSGGTAPSSFGSGOYEHYHIFDOMQOARAEDEAKWKREITGRGIPEROKG 480
DB 421 vkapflgsggtapssfgsgoyehyhifdomqoaradedeakwkreitgrgipergokg 480
QY 481 OLAVERAKOVVEEFLORKEAMONKARAGHMVYLARLROIRLONFNROOIRAKLGEKK 540
DB 481 olaverakovveeflorkeam onkaraghmvy larlrqlrqnfnroo iraklgekk 540
QY 541 EANHSEGCOSFEADMRKKITFSLKAHANARAVALKEQLERKREATEREKKWEEHLYA 600
DB 541 eanhsegcosfeadmrkkitfslkahanara valkeqlerkrereaterekkweehly 600
QY 601 KGVKSDVSPPLQGHETGSGPSKQOMRSVITSALKEVGVDSLTDTRETSEEMOKTN 660
DB 601 kgvksdvspplqghetgsgpskqomrsvits alk evgvds ldtretseemoktn 660
QY 661 AATSKREILRLNENIKAOEDEKGMQLSDTFEINVEDKKEHEKEKSVSSDRKKKEWAG 720
DB 661 aatskreilrlnenikaodek gmqnlsdtfe invedkkeh ekeksvssdrkkkewag 720
QY 721 QLVIPLDELTLDTSTSTERTVGVKILGNGSPRRAMGKSPDVSYLKILGAELQLOLQ 780
DB 721 qlvipldeltdtststertv gvki lgnsgprramgk spdvsylk ilgaelqlolq 780
QY 781 ELLENTTIRSETSPGEGYKPLINGEKKVOCISHEINPSAVIVOSPVTCKPESSEASPM 840
DB 781 ellenttirsetspgegykpl ingekk vocishe inpsavivospvtckp esseaspm 840

DB 781 ellenttirsetspegekykplltgtgkqvclshelpalvdspsvctkspefaeaspm 840
QY 841 SLKLEGNLEPPDDLEPFILOEPSTGNKDSLPCTTIDVWISSEKRTKETSADRTIOEN 900
DB 841 slklegnleppddlefi loepstgnkds lpc tti dvwi ssekrtk etsadrtioen 900
QY 901 EYSDGVSSVVDLSIDHITPGTNDOSHKCDVDSVOPEPFHKKVYSEHLNLYPOVQS 960
DB 901 eysdgvssvvdlsidhitpgt ndos hkc dvdsvo pefhkkv ysehl nlypovqs 960
QY 961 VQCSPEESFAFRSHSLPFRKNKNLIGLSTGFLDANPNPMLTCSLPLSLKFLRM 1020
DB 961 vqcspeesf af rshslpfrknkn l i g l s t g f l d a n p n p m l t c s l p l s l k f l r m 1020
QY 1021 DVPTVGDVRODNLLEIDEIKDENIKESPSDEDTVEETDTLOELQASMEQLLRBPQEE 1080
DB 1021 dvptvgdvrdnll eideik deni kesps de dtveetdt loelqas meqlr bpqee 1080
QY 1081 YSEEESEVLKNSDVEPTANGTDVADEDNPPSSBSALNEBMSDNGEIASCECDSVFN 1140
DB 1081 yseeeesv lkn s d v e p t a n g t d v a d e d n p p s s b s a l n e b m s d n g e i a s c e c d s v f n 1140
QY 1141 HLEELRLHLEBOEMGFEEFVEYKIKAIHEDEDENEICSKIVONILGNEHQHLYAKILH 1200
DB 1141 hleelrlhleboemgf eefve yekika i h e d e d e n e i c s k i v o n i l g n e h q h l y a k i l h 1200
QY 1201 LYMDAGAYORDNDE 1214
DB 1201 lymdagayordnde 1214

RESULT 2
AA39211
ID AA39211 standard; Protein; 1214 AA.
XX
AC AA39211:
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2356.
XX
KW Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; leukoemia; thrombolytic; drug screening; arthritis; inflammation;
KW
XX Homo sapiens.
XX
OS WO200153312-A1.
XX
PN 26-JUL-2001.
XX
PD 26-DEC-2000; 2000MC-US34263.
XX
PE 21-JAN-2000; 2000US-0488725.
XX
PR 23-APR-2000; 2000US-0532317.
XX
PR 09-JUL-2000; 2000US-0598042.
XX
PR 19-JUL-2000; 2000US-0620312.
XX
PR 03-AUG-2000; 2000US-0653450.
XX
PR 14-SEP-2000; 2000US-0662191.
XX
PR 19-OCT-2000; 2000US-0693036.
XX
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao QA, Zhou P, Goodrich R, Dmanac RT;
XX
XX WPI: 2001-442253/47.
XX N-PSDB: AA158367.

XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -

XX Example 4: SEQ ID NO 2356; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161365) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytosstatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/Inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

CC Sequence 1214 AA:

Query Match 99.9%; Score 6237; DB 22; Length 1214;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1213; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRYRLQKIGEGSKKALLIKSTEDGROYIKENINISRRSSRESEREAVAVLANMKH 60
 DB 1 MKRYRLQKIGEGSKKALLIKSTEDGROYIKENINISRRSSRESEREAVAVLANMKH 60
 QY 61 PRTIVYRSEFEENGLIYMDYCGGDLFRINAKGVLFQEQIILDMVQICLAKIYH 120
 DB 61 PRTIVYRSEFEENGLIYMDYCGGDLFRINAKGVLFQEQIILDMVQICLAKIYH 120
 QY 121 DKKILHARDIKSONIFLTQDGVQLDFGIARVLNLSVEIARCTIGTPYVLSPEICENKPY 180
 DB 121 DKKILHARDIKSONIFLTQDGVQLDFGIARVLNLSVEIARCTIGTPYVLSPEICENKPY 180
 QY 121 DKKILHARDIKSONIFLTQDGVQLDFGIARVLNLSVEIARCTIGTPYVLSPEICENKPY 180
 DB 121 DKKILHARDIKSONIFLTQDGVQLDFGIARVLNLSVEIARCTIGTPYVLSPEICENKPY 180
 QY 181 NNRKSDIMALGCVLYELCTLKHAFAGSMKNLVKLIISGSPPVSLHYSDLSRLVSQLEK 240
 DB 181 NNRKSDIMALGCVLYELCTLKHAFAGSMKNLVKLIISGSPPVSLHYSDLSRLVSQLEK 240
 QY 181 NNRKSDIMALGCVLYELCTLKHAFAGSMKNLVKLIISGSPPVSLHYSDLSRLVSQLEK 240
 DB 181 NNRKSDIMALGCVLYELCTLKHAFAGSMKNLVKLIISGSPPVSLHYSDLSRLVSQLEK 240
 QY 241 RNPDRPSPVNSILKEGFIAKRIEKFELSPQLAEFECLTKPSKSGSOPIAKPRASGNSI 300
 DB 241 RNPDRPSPVNSILKEGFIAKRIEKFELSPQLAEFECLTKPSKSGSOPIAKPRASGNSI 300
 QY 241 RNPDRPSPVNSILKEGFIAKRIEKFELSPQLAEFECLTKPSKSGSOPIAKPRASGNSI 300
 DB 241 RNPDRPSPVNSILKEGFIAKRIEKFELSPQLAEFECLTKPSKSGSOPIAKPRASGNSI 300
 QY 301 SYMPAQKITPKPAKYGIPLAKKYGGDKKLEHKKRPLQKHQAHOPEKRVNTGSEERKISE 360
 DB 301 SYMPAQKITPKPAKYGIPLAKKYGGDKKLEHKKRPLQKHQAHOPEKRVNTGSEERKISE 360
 QY 301 SYMPAQKITPKPAKYGIPLAKKYGGDKKLEHKKRPLQKHQAHOPEKRVNTGSEERKISE 360
 DB 301 SYMPAQKITPKPAKYGIPLAKKYGGDKKLEHKKRPLQKHQAHOPEKRVNTGSEERKISE 360
 QY 361 EAARPRLEPTEFEKKKKDOILSIMKAKOMKROKRELERINAROGMRNVLASAGSGE 420
 DB 361 EAARPRLEPTEFEKKKKDOILSIMKAKOMKROKRELERINAROGMRNVLASAGSGE 420
 QY 361 EAARPRLEPTEFEKKKKDOILSIMKAKOMKROKRELERINAROGMRNVLASAGSGE 420
 DB 361 EAARPRLEPTEFEKKKKDOILSIMKAKOMKROKRELERINAROGMRNVLASAGSGE 420
 QY 421 VAPFTLGSAGTIASSSESSRCQYEHYHAIFPOMQOAAENNEKMKREITGRILPRKOG 480
 DB 421 VAPFTLGSAGTIASSSESSRCQYEHYHAIFPOMQOAAENNEKMKREITGRILPRKOG 480
 QY 421 VAPFTLGSAGTIASSSESSRCQYEHYHAIFPOMQOAAENNEKMKREITGRILPRKOG 480
 DB 421 VAPFTLGSAGTIASSSESSRCQYEHYHAIFPOMQOAAENNEKMKREITGRILPRKOG 480
 QY 481 QLAVERAKOVEEFLQKREKADAMONKARAGHMYIARLQIIRLONFNRQOIKAKLNGEKK 540
 DB 481 QLAVERAKOVEEFLQKREKADAMONKARAGHMYIARLQIIRLONFNRQOIKAKLNGEKK 540
 QY 481 QLAVERAKOVEEFLQKREKADAMONKARAGHMYIARLQIIRLONFNRQOIKAKLNGEKK 540
 DB 481 QLAVERAKOVEEFLQKREKADAMONKARAGHMYIARLQIIRLONFNRQOIKAKLNGEKK 540
 QY 541 EANHSEGOEGSEADMRKKITLESIAHANARAVALKQLEKRRKEAVEREKVWEHLVA 600
 DB 541 EANHSEGOEGSEADMRKKITLESIAHANARAVALKQLEKRRKEAVEREKVWEHLVA 600
 QY 541 EANHSEGOEGSEADMRKKITLESIAHANARAVALKQLEKRRKEAVEREKVWEHLVA 600
 DB 541 EANHSEGOEGSEADMRKKITLESIAHANARAVALKQLEKRRKEAVEREKVWEHLVA 600
 QY 601 KGVKSDVSPPLGQHEHETGSPSKQOMRSVITSYALKVEGVDSLTDTRETSEMOKTNN 660
 DB 601 KGVKSDVSPPLGQHEHETGSPSKQOMRSVITSYALKVEGVDSLTDTRETSEMOKTNN 660
 QY 601 KGVKSDVSPPLGQHEHETGSPSKQOMRSVITSYALKVEGVDSLTDTRETSEMOKTNN 660
 DB 601 KGVKSDVSPPLGQHEHETGSPSKQOMRSVITSYALKVEGVDSLTDTRETSEMOKTNN 660

QY 661 AISKREILRLRNLNKAODEKQONTSDTFEINYHEDAKHEHEKSVSDRRKMEAGG 720
 DB 661 AISKREILRLRNLNKAODEKQONTSDTFEINYHEDAKHEHEKSVSDRRKMEAGG 720
 QY 721 QLVPLDELDTLDTSPSTERRHTVEVTKLGPNSPRANGKSPSPSVLKILGEALQLOT 780
 DB 721 QLVPLDELDTLDTSPSTERRHTVEVTKLGPNSPRANGKSPSPSVLKILGEALQLOT 780
 QY 781 ELLENTTIRSEIPEGGEKRYLITGEEKKQOCISHEINPSAIVDSVYEMKSPFSEASPM 840
 DB 781 ELLENTTIRSEIPEGGEKRYLITGEEKKQOCISHEINPSAIVDSVYEMKSPFSEASPM 840
 QY 841 SKLEGNLLEPPDLTEITLQEPSTNKDESILPCTITDVAIASEKETKQSDRTTODN 900
 DB 841 SKLEGNLLEPPDLTEITLQEPSTNKDESILPCTITDVAIASEKETKQSDRTTODN 900
 QY 901 EYSEDESVSTYDQSLDHIIEFGTNDQSHSKCDVKSVOPEPFHKKVYSEHLNLVPOYS 960
 DB 901 EYSEDESVSTYDQSLDHIIEFGTNDQSHSKCDVKSVOPEPFHKKVYSEHLNLVPOYS 960
 QY 961 VOCSPSESPFERSHSLPKNKNKNSLLIGLSTGLFDANNPKMLRTCSLPDLISKLFRILM 1020
 DB 961 VOCSPSESPFERSHSLPKNKNKNSLLIGLSTGLFDANNPKMLRTCSLPDLISKLFRILM 1020
 QY 1021 DVPYGVDRKQDLIEIDELKIDENIKESGSDSEDIYFEETDIDLOEQASMBOLLREQGE 1080
 DB 1021 DVPYGVDRKQDLIEIDELKIDENIKESGSDSEDIYFEETDIDLOEQASMBOLLREQGE 1080
 QY 1081 YSESESVYKNSDVEPTANGTVDVADEDDNPSESALNEEHRSDNSDGETASECECDVFN 1140
 DB 1081 YSESESVYKNSDVEPTANGTVDVADEDDNPSESALNEEHRSDNSDGETASECECDVFN 1140
 QY 1141 HIEEIRLHIEOMGEFEFEYERIKAIHEDEDENIEICSKTYONILGNHOLYAKILH 1200
 DB 1141 HIEEIRLHIEOMGEFEFEYERIKAIHEDEDENIEICSKTYONILGNHOLYAKILH 1200
 QY 1201 LWADGAYQEDNDE 1214
 DB 1201 LWADGAYQEDNDE 1214

RESULT 3
 AAM39210
 ID AAM39210 standard; Protein: 1242 AA.
 XX AAM39210;
 AC 22-OCT-2001 (first entry)
 DT 22-OCT-2001 (first entry)
 DE Human polypeptide SEQ ID NO 2355.
 KW Human; nootropic; immunosuppressant; cytosstatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 OS Homo sapiens.
 XX Homo sapiens.
 PN WO200153312-A1.
 DB 26-JUL-2001.
 XX 26-DEC-2000; 2000MO-US34263.
 XX 21-JAN-2000; 2000US-0488725.
 XX 25-APR-2000; 2000US-0532317.
 XX 09-JUL-2000; 2000US-0596042.
 XX 19-JUL-2000; 2000US-0620312.
 XX 03-AUG-2000; 2000US-0653450.
 XX 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.

(HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA, Zhou P, Goodrich R, Drmanac RT;

DR N-PSDB; AA158366.

Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries -

PS Example 4; SEQ ID NO 2355; 10078pp; English.

CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM36642-AA442213) with neurotrophic,
CC immunosuppressant and cyostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, hemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, Leukemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

SQ Sequence 1242 AA;

Query Match	99.58;	Score 6213;	DB 22;	Length 1242;
Best Local Similarity	97.78;	Pred. No. 0;		
Matches 1213;	Conservative	0;	Mismatches 1;	Indels 28;
				Gaps 1

Qy	1	MEKVVRLKOTIGESGSKMLLVKSPEDGKOVYIEINISRMSSKERESREAVYLANMKH	60
Dp	1	mekyvrllqtlgsgsfykallvksctedgrvyikelnstsmstkeereavlanmkh	60
Qy	61	PNIVQRESFEENGSLIYIMDYCEGDLFRINAKGVLFQEOQLIDMFVQICALKHYH	120
Dp	61	pnlvqyresfeengsgslylwmdyceegdlfrlnakgvylfdeqqlidwlvqclalkhvh	120
Qy	121	DRKILHRDIKSONIFLTKRDGVQLGDGFARVINSTVELARTICGTPYLLSPETCEKNPY	180
Dp	121	drkllhrtdlksqnifltkrdgvtqldgdfarvlnstvelartictgpyllspetceknpy	180
Qy	181	NNKSDIWMALGCVLYELCTLKAFEGSGMKNLVKTIISGSPPYSLHYSDRLSYLSOLF	240
Dp	181	nnskdiwalgcvlyelctclkafeegsmknvlkllsgspshysdrlsydlrslvsqflk	240
Qy	241	RNRDPSVNSLIEGFLAKRIEKLPSPOLIAEEFLCTFKFSQSPPIPAKRASGQNSI	300
Dp	241	rnrddpsvnslliegflakrliekflspqliaeeftcltksqspipakrpsaqnsi	300
Qy	301	SVMPACKITKPAKYGIPLAYKKYGDKKLHEKKRPLQKHKQAOITPEKRVNTGEENRIISE	360
Dp	301	svmpackltckpaakygiplaykkygdckllhekkrplokhnqaoitpekrvntgeeenriise	360
Qy	361	EAARRRLEFIEKKOKDQIISLMKAEOMKROEKELERINARQGRNVLSSAGSGE	420
Dp	361	eaarrtlefiekkkqdqiiislmkegmkrqekelerlnatreqgrnvlssaagsge	420
Qy	421	VKAPFLGSGGTLAPSPSSRQGYEHYHATIDMOQOARADENAKKRELYRGCLPERQK	480
Dp	421	vkapflgsggtlapssfsrsgyehyhaltdmqgqaadeneakwrtelyrgclperqg	480

OY	481	QLEPRKQVVEFLPQRRREMAOKKAAEGM-----V	512
Db	481	qlaverakvyeefqrrreemqpkataeglmjllqnaamygrrpsarsgkprkveev	540
OY	513	YLAIRLQRIQLONFNEORQIKAKIRGKKRPNHSGEGSEEDMRRKIKESIAKANRA	572
Db	541	YLAIRLQRIQLNTERGQIKAKIRGKKRNHSGEGSEEDMRRKIKESIAKANRA	600
OY	573	AYLKQLEPRKKEAYEPEKKVWEHLVAKWSSDVPPLQGHFEGSPKQQRNRSYV	632
Db	601	avlkqleprkkrkeeyekkvweehlvakvksdvpplqghfegspskqqrmsvlysv	660
OY	633	TSALKEGVDSLLDTRETSSEMOKTNNAISSKREILRLNLNAKOEDKQMONLSDF	692
Db	661	tsalkegyvdslldtretseemqktnnaisskreilrlnlnlkaedeqyukqlldstfi	720
OY	693	EINNHEDAKKEHEKESVSSORKKRWKMGQGLVPLDELTLDTSFSTTERRHYVEYKILAPN	752
Db	721	einhedakehekeksvsdrkkwkgqglvpldeltdltdtsfsterttyevyklqpn	780
OY	753	GSPRRANGKSPSTDVLELILDEAELOLOTELLENTTTRSEISPEGRKRYPLITGEKKVCI	812
Db	781	gsprrangksptdsvlklilgeaelqlqtelentltrseispegekykpllitgekkvci	840
OY	813	SHEINPSAIVDSVPETKSPFSEASPOMSIKLGNLEPPDDETETLOPESGTNDESLP	872
Db	841	sheinpsaivdspetkspfseaspmsiklegnlleepddletellqepsgtndeisl	900
OY	873	CTITDVMSESEKKEKFNQSDRITIOENESYEDGVSSTVDOLSDNIHEPGTNDOSKCD	932
Db	901	ctitdvmseeketkeqsdritlqenevsegyvstvdqldslhiepntndgshkcd	960
OY	933	VDSVQRPFRPKKVVHSEHNLVPOVOSVCSSEESFAFRSHSHLPKNNKNSLILGLS	992
Db	961	vdsvqrpfrfkvhvshelnlvpqvsgvcspeesfattrshshlpknnknsllglsls	1020
OY	993	TGLEDANPKMLTRCSPDLSTKFLRLMVPYGVDRONLEIDELKQENIKEGSPDSE	1052
Db	1021	tgledanpkmrltrcspldstklflrlmvdvplvgdvrgnleldelkqenlkepsdse	1080
OY	1053	IVFEETPTDLQELQASNEQLRQPGEEVSEEESTLAKSDVPEPRANGTDVADEDNPS	1112
Db	1081	ivfeetdtdlqelqasneqllrtqpegveeseeestlaksdvpeprangtdvadeddnps	1140
OY	1113	ESALNEHWSNDSDGEIASFCECDVSFNLHELRLLEDEMKFEKFEVEYEKIKAIHDE	1172
Db	1141	esalneehwsnsgelasscedsvfnhleeirlhleqemfekkfevyekikalihede	1200
OY	1173	DENIECSKIYONITLGNHGHLYAKTILHLVMADGAYQEDNDE	1214
Db	1201	deniecskiyqnligneqhlyaktilhlvmadgayqednde	1242
RESULT 4			
ID	AAU07103	standard; Protein; 1007 AA.	
AC	AAU07103:		
DF	24-Oct-2001	(first entry)	
XX	XX	Human novel human protein, NHP #3.	
XX	XX		
KW	XX	Human: novel human protein: NHP. breast cancer: prostate cancer:	
XX	XX	immunogen: antibody; gene therapy; antisense.	
OS	XX	Homo sapiens.	
XX	XX		
XX	XX	WO200161016-A2.	
PD	23-AUG-2001.		
XX	XX		
XX	15-FEB-2001; 2001WO-US05356.		


```

XX 18-FEB-2000; 2000US-0183582.
PR 22-FEB-2000; 2000US-0184014.
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Walke DW, Hu Y, Nepomichy B, Turner CA, Zambrowicz B:
XX WPI: 2001-502793/55.
XX N-PSDB: AAS11559.
XX
XX Isolated nucleic acids encoding novel human proteins useful for the
XX treatment of disease and as probes for testing and detection
XX
XX Claim 5; Page 41-43; 69pp; English.
XX
XX The invention relates to novel human proteins (NHP) and the nucleic
XX acids encoding them. The nucleic acids encode mammalian transporter
XX proteins and are useful for the treatment (e.g. by gene therapy or
XX antisense technology) of any of a wide variety of symptoms associated
XX with biological disorders (e.g. breast and prostate cancer) or imbalances
XX and as probes for the identification, selection and validation of novel
XX molecular targets for drug discovery. The proteins may be used to raise
XX anti-NHP antibodies. The present sequence represents an NHP of the
XX invention.
XX
XX Sequence 1007 AA:
SQ
Query Match 82.7%; Score 5162; DB 23; Length 1007;
Best Local Similarity 100.0%; Pred. No. 2,1e-303;
Matches 1007; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 208 MNKLVKTISSGPPVSLHRSYDRLSVQLFRNPRDRSPVNSILEKGFIAKRIEKLPLS 267
DB 1 mknlvkltissgppvslhysydlrslvsglfrnprdrpsvnsilekglfakrleklis 60
QY 268 POLIAEPCLTFSKFSQPIPAKRPASGNSISVMPAOKITIPAKYGIPLAYKKYDK 327
DB 61 plliaeefcltfskfsqpiipakrpasgnsisvmpaqitkpaakygiplaykkydk 120
QY 328 KLHEKKPLOKHOAHQTPPEKRVNTGEERRKISEBAARRRLEFIEKEKKQKDOIISLKA 387
DB 121 klhekkplokhoahqtppekrvntgeerrkiseeaarkrrlefeiekekqkdqllslka 180
QY 388 EGMKROEKEERLERINRAEOGMRNVLSAGSGEYKAPFLSGGTIAPSSFSRGOYENH 447
DB 181 egmkroekelerinraeogmrnvlsagsgyevkapflsggtiapsfsfsgyenyh 240
QY 448 AIFDMOOORADENAKKREIYGRGLPEROKGOLAVERAQYEEFLQKRRAMONKAKA 507
DB 241 aifdmqgqreedneakkreiygrglperqkgqlaverakyeeflqkrramqkaka 300
QY 508 EGHMYIARLRIORIONENROOIKAKLKGKKRANHSEGOSEEDMRKKRIESTLKAH 567
DB 301 eghmyiarlriqirgnfnerqgkakilrgekkranhseggseeadmrirkkieslka 360
QY 568 ANAKRAVLKEOLEERKREKAYEREKKWEHLVAKGVSSDVSPPLOHETGSGSPKQOMR 627
DB 361 anaraavlkeqlerkrkayerekkwehnlvakgvssdvspplqhetgsgspkqomr 420
QY 628 SVIYSTALKEVGVDSLTITRETSSEMOKTNNAISSKRILKRLNENLKAODEKGMN 687
DB 421 svlsvtsalkevvdsltdtreseemqknnaiskrrilrlnenlkaodekgmn 480
QY 688 LSDTEFINVEDAKHEKEKSVSDRKKWEAGGOLVLPDELDTJDSFTTERHVGAYI 747
DB 481 lsdtefinvhedakehekevssdrkkweagglvlpdeltdjdsfsterhvgayl 540
QY 748 KLGNGSPRRANGKSPFSDVYLKILGAEALQLOTLELBNNTIRSEISPEGEKYPPLITGK 807
DB 541 klgnpsprrangksptfsvylkilgaealqlotlelbnntirseispegekypplltgk 600

```

```

QY 808 KVOCISHINPSAIVDSPVETKSPFEFSASPMQSLKLEGNLEEPDLETTILOEPSGNTK 867
DB 601 kvocishinpsaivdspvetkspfefsaspmqslklegnleepdletteilqpsgntk 660
QY 868 DESLPCTITDVAWIESEKETKETOSADRITIOENREVSDEDSVTSVQOLSDHIEGTDNQ 927
DB 661 deslpctitdwaiseeketketgsadrilcigenevsedgsvstvdqsdshiepgtdnsq 720
QY 928 HSKCDVKSVPQEPFEPFKVYHSEHNLVPOVOSVOCSPSEFAFRSHLPPKKNKNSL 987
DB 721 hskcdvksvqepfepfkvyhsehlvpgvsvqcspeestafstshlppkknksl 780
QY 988 LIGSTGFPDANNKMLRTCSLPDLKIRPTLMDPYGVGDYRQDMLTDEIKDENIKGP 1047
DB 781 ligstgldannpkmrlrtcsipolskirltmdpvcvgdvqdaleldelkdenikegp 840
QY 1048 SDESDIVEETDLDLOELQASMEQLREOPGEYSSEESVYKNSDVEPTANGTDVAD 1107
DB 841 sdsediveetdldlqelqasmeqlrqpgeesyseesvlyknsdveptangtdaved 900
QY 1108 DNPSESALNEEWHSDNSDEIASSECECDVFNHLELRLHLEQMGFEKFEYETIKA 1167
DB 901 dnpsesalneewhshdsdelscecdsvfnhlelrlhleqmgfeyekyeka 960
QY 1168 IHEDDENIEICSKIVONILGNEHQHLYAKITLHVADGAYOEDNDE 1214
DB 961 ihedenielciskivgnilgnehqhlyakitlhvmadgayednde 1007

RESULT 5
AAU07101
ID AAU07101 standard; Protein; 1035 AA.
XX
XX AAU07101;
XX
XX 24-OCT-2001 (first entry)
XX
XX Human novel human protein, NHP #1.
XX
XX Human; novel human protein; NHP; breast cancer; prostate cancer;
XX immunogen; antibody; gene therapy; antisense.
XX
XX Homo sapiens.
XX
XX WO200161016-A2.
XX
XX 23-AUG-2001.
XX
XX 15-FEB-2001; 2001WO-US05356.
XX
XX 18-FEB-2000; 2000US-0183582.
XX
XX 22-FEB-2000; 2000US-0184014.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Walke DW, Hu Y, Nepomichy B, Turner CA, Zambrowicz B:
XX WPI: 2001-502793/55.
XX
XX Isolated nucleic acids encoding novel human proteins useful for the
XX treatment of disease and as probes for testing and detection
XX
XX Claim 2; Page 33-35; 69pp; English.
XX
XX The invention relates to novel human proteins (NHP) and the nucleic
XX acids encoding them. The nucleic acids encode mammalian transporter
XX proteins and are useful for the treatment (e.g. by gene therapy or
XX antisense technology) of any of a wide variety of symptoms associated
XX with biological disorders (e.g. breast and prostate cancer) or imbalances
XX and as probes for the identification, selection and validation of novel
XX molecular targets for drug discovery. The proteins may be used to raise
XX anti-NHP antibodies. The present sequence represents an NHP of the
XX invention.

```

xx Sequence 1035 AA:

Query Match 82.3%; Score 5138; DB 22; Length 1035;
Best Local Similarity 97.3%; Pred. No. 6e-302; 0; Indels 28; Gaps 1;
Matches 1007; Conservative 0; Mismatches 0;

```

OY 208 MKNVYKTIISGSPVSLHYSYDLSLVSLQFRNPRDRSVSLLEKGFIAKRIKFLS 267
DB 1 MNLVKIVISGIPPSVSLHYSYDLSLVSLQFRNPRDRSVSLLEKGFIAKRIKFLS 60
OY 268 POLIAEEFCILKTSKESGQIPAKRPAASGONSISVMPAKITKPAKYGIPLAYKKYGDK 327
DB 61 PQLIAEEFCILKTSKESGQIPAKRPAASGONSISVMPAKITKPAKYGIPLAYKKYGDK 120
OY 328 KLHEKKPLQKHQAHOQTEPKRVNTGEERRKISEAARKRRLEPIEKEKKOKDOISLMA 387
DB 121 KLHEKKPLQKHQAHOQTEPKRVNTGEERRKISEAARKRRLEPIEKEKKOKDOISLMA 180
OY 388 EOMKROEKERLERINRAREQGMWNVLSAGSGEYKAPFLSGGTTIAPSSFSRSGOYEHTH 447
DB 181 EGMKRGKERTLERINRAREQGMWNVLSAGSGEYKAPFLSGGTTIAPSSFSRSGOYEHTH 240
OY 448 AIFDQMOOQRAEDNEAKMKRREIYRGLEPEROKOLAVERAKOYEETLORRREAMONKARA 507
DB 241 AIFDQMGQRAEDNEAKMKRREIYRGLEPEROKOLAVERAKOYEETLORRREAMONKARA 300
OY 508 EGHM-----VYIARLROIIONNEBOOIKAKIRGEX 539
DB 301 EGHM-----VYIARLROIIONNEBOOIKAKIRGEX 360
OY 540 KEANHSEGESEADMRKRIKIESLKHANARAVLKEOLERKRAKAYEREKWEHELTV 599
DB 361 KEANHSEGESEADMRKRIKIESLKHANARAVLKEOLERKRAKAYEREKWEHELTV 420
OY 600 AKGVKSSVSPPLCOHETGSPSKQOMRYSIVTSALKEVGVDSLTDTRETSEEMOKTN 659
DB 421 AKGVKSSVSPPLQGHETGSPSKQOMRYSIVTSALKEVGVDSLTDTRETSEEMOKTN 480
OY 660 NAISSKREILRLNENLKAODEKGMONLSDTREIYNVEDAKHEKESVSDRKKWEAG 719
DB 481 NAISSKREILRLNENLKAODEKGMONLSDTREIYNVEDAKHEKESVSDRKKWEAG 540
OY 720 GOLVPIDELTLDTSFSTTEHNTYGEVIRKLGPNSSPRAMGKSPTOVSUKILGEAEILQ 779
DB 541 GOLVPIDELTLDTSFSTTEHNTYGEVIRKLGPNSSPRAMGKSPTOVSUKILGEAEILQ 600
OY 760 TELLENTTIRSEISPEGEKRYPLITGEKKYOCISHEINPSAIVDSVVERKSPFESASPO 839
DB 601 TELLENTTIRSEISPEGEKRYPLITGEKKYOCISHEINPSAIVDSVVERKSPFESASPO 660
OY 840 MSIKIEGTLPEPDLETEILOEPSTGNKDSLCCTTDWISEREKTKEQOSADRTIOE 899
DB 661 MSIKIEGTLPEPDLETEILOEPSTGNKDSLCCTTDWISEREKTKEQOSADRTIOE 720
OY 900 NEVSDEGVSTVVDOLSDIHEPGTNDOSKCDVDSVOPPEPFKRVHSEHLNLTVPQ 959
DB 721 NEVSDEGVSTVVDOLSDIHEPGTNDOSKCDVDSVOPPEPFKRVHSEHLNLTVPQ 780
OY 960 SVOCSPESFAFRSHSLPRKNNKNSLIGLSTGLFDANPMRLTCSLPDLISKILFRL 1019
DB 781 SVOCSPESFAFRSHSLPRKNNKNSLIGLSTGLFDANPMRLTCSLPDLISKILFRL 840
OY 1020 MDVPVGVGDVNDNEIDEIKENIKESPSSEDIIVEETDTDOEQASMEQLLRPOPE 1079
DB 841 MDVPVGVGDVNDNEIDEIKENIKESPSSEDIIVEETDTDOEQASMEQLLRPOPE 900
OY 1080 EYSEEESEVILKNSVPEPANTDVADBDNPPSSASALNEMWHSNDNSGATASCECDVSF 1139
DB 901 EYSEEESEVILKNSVPEPANTDVADBDNPPSSASALNEMWHSNDNSGATASCECDVSF 960
OY 1140 NHLEELRLHLEQMGFEKFEVYETKAIKIHEDDENIEICSKIVONILGNHOHLTKIL 1199

```

```

DB 961 NHLEELRLHLEQMGFEKFEVYETKAIKIHEDDENIEICSKIVONILGNHOHLTKIL 1020
OY 1200 HLVMADGAYOEDNDE 1214
DB 1021 HLVMADGAYOEDNDE 1035

RESULT 6
AAM40996
ID AAM40996 standard; Protein: 966 AA.
AC AAM40996;
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 5927.
XX
KW Human; nootropic; immunosuppressant; cytosolic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN W0200153312-A1.
PD 26-JUL-2001.
PE 26-DEC-2000; 2000MO-US34263.
PF
PR 21-JAN-2000; 2000US-0488725.
PR 23-APR-2000; 2000US-0532317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Trang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR N-PSDB; AAI60152.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries.
PS
XX
XX Example 2: SEQ ID NO 5927; 10078pp: English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytosolic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localized neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: immune system suppression,
XX activin/inhibin activity, chemotactic/chemokine activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.

```

SO	Sequence	966 AA:	
	Query Match	75.98; Score 4735.5; DB 22; Length 966;	
	Best Local Similarity	96.38; Pred. No. 1.1e-277;	
	Matches 930; Conservative	2; Mismatches 5; Indels 29; Gaps 2;	
QY	278	KTESKFGSOP1-PAKRASGONSIWVPAOKITTKPAKYGPPLAVKYYGDKRLHEKKPLQ	336
DB	1	khiskfsgalylq1ktpasqgniswmpakltpakkyg1plakkygdkllhekkplq	60
QY	337	KHKQAHOTPEKRYNTGEEERKRISEEAARRRLFEIEKEKKOKDQIITSLMKAEOMKROKE	396
DB	61	khkqahotpekryntgeerksiseeaarrlrlelekekkqgdqllsmkaeqmrgqke	120
QY	397	RLERINBAROGWRNVLSAGSGEVKAPPLGSGGTAPSSFSRQCYEHYAIRPQMOQO	456
DB	121	rlerinbarogwrvnlvsagsgdevkapflgsggtlapssfsrsgyhyairfdmqgq	180
QY	457	RAEDNEAKMKREIYGRQLPEROKGOLAVERAKOVEEFLORREKREAMONKARABGHM----	511
DB	181	raedneakmkreiygrqlperqgqlaverakqveeflgrcreamqnararegmglqgn	240
QY	512	-----VYLARLROIRLONFNPNROIRKALKEGKREKRNHSGO	548
DB	241	laamygrpsstrgqkprnkeevyarlrlrqlqinergqlkallgekkkaanhsegq	300
QY	549	EGSEADMRKKKTIESTLKAHANARAVALKEOLERKKREAYEREKKVWEHLVAKGVKSSDV	608
DB	301	egseadamrkkkleslkanaharaavalkeqlerkrkeayerekkvwehlvakgvkssdv	360
QY	609	SPLGHEGTGSPSKOQMSVSVTSALKVEGVDSLTDRTRSEMOKTNNATISSKREI	668
DB	361	splghegtgspskqmsvsvtsalkvegvdsllttrtseemkktlnaalskrel	420
QY	669	LRLNENLKAQOEDEKGMQLSDTFEINVEDAKHEHEKESVSSDRKKWEAGQOLVPLDE	728
DB	421	lrlnenlkaqodekqqlsdtfelnvedakehekesvsdsdrkkweagqlvplde	480
QY	729	LTLDTFSFTEERTVEEVILGPNCSPPRAMGKSPDVSALKLIGAEVQLTTELEMTTI	788
DB	481	ltdltsfsterltvgevlilgpnsprrawgkspdvylkljgaeeilqtellemtti	540
QY	789	RSEISPEGEKYPPLITGEKKVOCISHEINPSAIVDSPVTKSPFSEASPMOSIKLEGNL	848
DB	541	rseispegekypplltgekkvgqlsheinpsaivdspvtskpfseaspmosiklagnl	600
QY	849	EEPDLTEELLQPSGTNNKDESLPCTITDVMISEEKERKQMSADRITQENEVSEQVGS	908
DB	601	eeppdlteellqpsgtlnkdeslpcctltvwiiseeketkeqsaadrilqeneyseqgvs	660
QY	909	STVQSLDHIIEPGTNDSDHSCDVNKSVPPEPFHKKVYHSHLVLVQOVSVQCSPEPS	968
DB	661	stvqslldhiiepgtndsgnscdvksvqpepfhkvynshlnlvpqvgvscpsps	720
QY	969	FAFSSHSLPPKNNKNSLLIGLSTGLPANNPKMLRTCSLPLDLSKLFRTLMADVPVGDV	1028
DB	721	fafsshslppknnknslllglstglidampkmlrtcslpdlsklrtlmadvplvgdv	780
QY	1029	RQDNLEIDEIKDENIKEGPSDESIVFEETDIDLOELQASMEQLLREQPGEEYSSEERSV	1088
DB	781	rqdnleideidenikegpsdesivfeetdldelqasmeqllrepgseyseesv	840
QY	1089	LKNSDVEPTANGTDAVEDDNPSSSALNEMHSDNSDGETLASECCDSVNHHEERLH	1148
DB	841	lknsdveptangtcdvadednpssesalnemhshsdngelaseccdsvfnhheerlh	900
QY	1149	LEQSMGFKEFFEYERIKAIHEDEDENIEICSKIYONITIGNHQLYAKIILHVAWAGAY	1208
DB	901	legsmgfekfeyekrikaihededenielskivgnllgnehqllyaklilhvmadgay	960
QY	1209	QEONDE 1214	

DB	961	gednde 966	
	RESULT 7		
	AA040997	standard; Protein; 966 AA.	
ID	AA040997		
XX	AA040997;		
DT	22-OCT-2001	(first entry)	
XX			
DE	Human polypeptide SEQ ID NO 5928.		
KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;		
KW	chemoknetic; thrombolytic; drug screening; arthritis; inflammation;		
KW	leukaemia.		
OS	Homo sapiens.		
PN	WO200153312-A1.		
PD	26-JUL-2001.		
XX			
PF	26-DEC-2000; 2000WO-US34263.		
XX			
PR	21-JAN-2000; 2000US-0488725.		
PR	25-APR-2000; 2000US-0552317.		
PR	09-JUL-2000; 2000US-0598042.		
PR	19-JUL-2000; 2000US-0620312.		
PR	03-AUG-2000; 2000US-0653450.		
PR	14-SEP-2000; 2000US-0662191.		
PR	19-OCT-2000; 2000US-0693036.		
PR	29-NOV-2000; 2000US-0727344.		
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
PI	Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;		
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;		
PI	Zhao QN, Zhou P, Goodrich R, Drmanac RT;		
XX			
DR	WPI: 2001-442253/47.		
XX			
DR	N-PSDB: AA160153.		
XX			
PT	Novel nucleic acids and polypeptides, useful for treating disorders		
PT	such as central nervous system injuries -		
XX			
PS	Example 2; SEQ ID NO 5928; 10078bp; English.		
XX			
CC	The invention relates to human nucleic acids (AA157798-AA161369) and		
CC	the encoded polypeptides (AA038642-AA042213) with nootropic,		
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful		
CC	in gene therapy. A composition containing a polypeptide or polynucleotide		
CC	of the invention may be used to treat diseases of the peripheral nervous		
CC	system, such as peripheral nervous injuries, peripheral neuropathy and		
CC	centralised neuropathies and central nervous system diseases, such as		
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic		
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the		
CC	utilisation of the activities such as: Immune system suppression,		
CC	Activin/inhibin activity, chemotactic/chemoknetic activity, haemostatic		
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,		
CC	assays for receptor activity, arthritis and inflammation, leukaemias and		
CC	C.N.S. disorders.		
CC	Note: The sequence data for this patent did not form part of the printed		
CC	specification.		
XX			
SO	Sequence	966 AA:	
	Query Match	75.98; Score 4735.5; DB 22; Length 966;	
	Best Local Similarity	96.38; Pred. No. 1.1e-277;	

Matches	930: Conservative	2: Mismatches	5: Indels	29: Gaps	2:
OY	278	KTESKFGSGPI-PAKRPASGONSISVPAQKITTAKPAKGIPLAVKAKKGGDKLHEKRPLO	336		
Db	1	khfskfsgqalylgltkrpasgqnsiswmpagkltkpaakyglpkygdkllhekkpqlg	60		
OY	337	KHKQAOOTPEKRYNTEGEERKRISEEAARKRLRLEPIEKEKKOKDOIISLMKAEOMKROKE	396		
Db	61	khqgahqtpkryntegeerrkiseeaarrrtleflekkygdqllismkgeqmkrygke	120		
OY	397	RLEKRNARPOGWRNVLSCGSEVAPFLGSGGTAPSSFSRCROYEHYNAIPQOQO	456		
Db	121	rletlnareqgwrnvlsgsgsevakpflsgsgtlapsstisryqeyhahldqmqgq	180		
OY	457	RAEDNEAKWREIYGRQLPEBROKOLAVERAKOVEBFLORREKRAMONKARAGHW----	511		
Db	181	raedneakwkrelygrqlperqqlaverakqveeflqkrreamqnkaregllmq	240		
OY	512	-----VYLARLRQIRLQNFNERQOIKAKLGEKKEKANHSEGO	548		
Db	241	laamygrpsrsrgdkprnkseevylarlrlqlqnfnerqqlaklrgkekanehseqg	300		
OY	549	EGSEADMRKRIESTLKAHANARAVALKEQLERKREKREKWEELVAKGVKSSDV	608		
Db	301	egseeadmrkrkieslkahanaraavlkeqlerkrerekkvweelhvakgykssdv	360		
OY	609	SPPLGHEHTGSGSKOOMSIVTSALKEVGVDSTLTPTRETSPEMKTNNATSKREI	668		
Db	361	spplghehtgsgskqgmsvtsalkevyvdsaltdtretseemqktmnaisskrei	420		
OY	669	LRLNEMTKROEDEKMONSDPFETNVEDAKENHEKESVSDPKKKEAGOLYIPDE	728		
Db	421	lrlnenlkegedekqnlstfelnvedakehekessvdsqkkaagqqlipde	480		
OY	729	LTLDTSTTERTVGEVILKLPNGSPRAKGSPPDVLKLGAEALQLOTELENTTI	788		
Db	481	ltdltssttertlvgevlklpngsprtrawksprdsylklgaaelqltelenttli	540		
OY	789	RSEISPEGEKYKPLITGEKKVOCISHEINPSAIVDSVETKSPESSEASPOMSLKEGSL	848		
Db	541	rseispegekykplltgekkvqclshelnpsalvdspevkspetsespsqmslkllegsl	600		
OY	849	EEPDDLETELLOPESGTNKNDESIPCTITDVIWISSEKREKTOGASDRITIOENEVSEDOVS	908		
Db	601	eeppddleltellpgegtlnkdesipctitdvwisseeketgsadritltqenevseodgvs	660		
OY	909	STVDQLSDHIEPGTNDQSHKCDVDKSVOPFFFKVYVSHENLIVPOVOSVQCSPEES	968		
Db	661	stvdqldshiepgtndsqhskcdvdksvqpefffkvvshenlilvpqvgvqcspees	720		
OY	969	FAFSSHRLPPKNNKNSLLIGSLGTFDANNPKMLRTCSLPDLKFLRFLMDVPTVDV	1028		
Db	721	fafsshrlppknnknslllgslgltfdannpkmlrtcslpdlkflrltmdvptvgdv	780		
OY	1029	RONMLRTDKDENIKEGSDSDIYFEETDIDLOELQASMOQLRBOGGEYXSEESV	1088		
Db	781	rgonmlrtdekenikegdsdsdiyfeetdldloelqasmqlrboqgeyseeesv	840		
OY	1089	LKNSDVPTANGTVADDEDNPSSEALAEKNSDNOCELASSECDSVFNHLEELRLH	1148		
Db	841	lknsdvptangtvadeddnpssealaeknsdndsgelaseecdsavfnhleelrlh	900		
OY	1149	LEDEMGFEKFEVEYERIKAIHEDEDENIEGSKIVONTIIGNHQHLYAKLHLMVADGAY	1208		
Db	901	legemgfekfeveyekaihededenietcskivgnliqnehqhlyaklhlhvmadgay	960		
OY	1209	QEDNDE 1214			
Db	961	gednde 966			

RESULT 8
AAV76753

ID	AAV76753 standard; Protein: 345 AA.
XX	
AC	AAV76753;
XX	
DT	17-APR-2000 (first entry)
XX	
DE	Human protein kinase homologue, PKH-6.
XX	
KW	Protein kinase homologue: human; PKH: diagnosis: therapy: cancer: AIDS:
KW	autoimmune disorder: inflammatory disorder: reproductive defect: asthma:
KW	diabetes mellitus; infertility; ovulatory defect; endometriosis;
XX	polycystic ovary syndrome.
OS	Homo sapiens.
XX	
PN	US6013455-A.
XX	
PD	11-JAN-2000.
XX	
PF	15-OCT-1998: 98US-0173581.
XX	
PR	15-OCT-1998: 98US-0173581.
XX	
PA	(INCY-) INCYTE PHARM INC.
PI	Hillman JL, Yue H, Yang YT, Corley NC, Gorgone GA, Azimzai Y;
PI	Lu DAM, Bandman O, Guegler KJ;
DR	N-PSDB: A4686797.
XX	
PT	Nucleic acids encoding a human protein kinase homolog useful for
PT	preventing, diagnosing and treating cancer, autoimmune/inflammatory
PT	disorders and reproductive defects -
PS	Claim 1: Column 51-54; 38pp: English.
XX	
CC	This sequence represents a human protein kinase homolog (PKH) of the
CC	invention. The PKH sequences may be used in the prevention, treatment and
CC	diagnosis of diseases associated with inappropriate PKH expression such
CC	as cancers, autoimmune/inflammatory disorders and reproductive defects.
CC	They may be used to treat disorders associated with decreased PKH
CC	expression such as cancers (e.g. lymphoma, melanoma and cancers of the
CC	breast lung and prostate), autoimmune/inflammatory disorders
CC	(e.g. AIDS, asthma and diabetes mellitus), and reproductive
CC	defects (e.g. infertility, ovulatory defects, endometriosis and
CC	polycystic ovary syndrome). The DNA may be administered to treat diseases
CC	by rectifying mutations or deletions in a patient's genome that affect
CC	the activity of PKH by expressing inactive proteins or to supplement the
CC	patients own production of PKH polypeptides. Additionally, the DNA may be
CC	used to produce PKH, according to standard recombinant DNA methodology,
CC	by inserting the nucleic acids into a host cell and culturing the cell to
CC	express the protein. Conversely, antisense nucleic acid molecules may be
CC	administered to down regulate PKH expression by binding with the cells
CC	own PKH genes and preventing their expression by DNA and antisense
CC	sequences may also be used as DNA probes in diagnostic assays to detect
CC	and quantitate the presence of similar nucleic acid sequences in samples,
CC	and hence which patients may be in need of restorative therapy. They may
CC	also be used to study the expression and function of PKH polypeptides and
CC	their role in metabolism. The PKH polypeptides may be used as antigens in
CC	the production of antibodies against PKH and in assays to identify
CC	modulators (agonists and antagonists) of PKH expression and activity. The
CC	anti-PKH antibodies and PKH antagonists may also be used to down regulate
CC	PKH expression and activity. The anti-PKH antibodies may also be used as
CC	diagnostic agents for detecting the presence of PKH polypeptides in
XX	samples.
XX	
SO	Sequence 345 AA:

Query Match 27.8%; Score 1734; DB 21; Length 345;
Best Local Similarity 90.5%; Pred. No. 6.2e-97;
Matches 341; Conservative 3; Mismatches 1; Indels 32; Gaps 1;

```

OY 1 MEKYVRLQKIGESFGKAILVSTEDGROYVKEINISMSKEREESREAVAVLANMKH 60
    |||||||
DB 1 mekyvrlqkigsgfkgallvkstcdgryvkeinismskereeerravavlanmkh 60
OY 61 PNIVQYRESFENGSLYVMDYCEGDLFKRINAOKGVLFQEDQILDFVQICLAKHVN 120
    |||||||
DB 61 pniyqyresfe-----gllwdfvqiclaikhvn 88
OY 121 DRKILHRDIKSONIFLTRDGTVOLGDFGIARVNSTVELARCTICGTPYLSPEICEKMPY 180
    |||||||
DB 89 drkllhrdiksqnifltrdgtvlgdfgiarvlnstvelarctcigtpylspceicemkpy 148
OY 181 NNSKDIWALGCVLYELCTLKHAFAGSMKNLVKTIISGSPPVSLHYSYDLRSLVSOLF 240
    |||||||
DB 149 nksdiwalgcvlyelctlkhafeagsmknlvklisgsfpvslhysydlrslvsqflk 208
OY 241 RNPDRPSVNSILEKGFIAKRIEFLSPQILAEFCLTKFSKFSQPIPARKRPASGNSI 300
    |||||||
DB 209 rnpdrpsvnsilekgfiakrleflspqilaefcltkfskfsqpiarkrpsagqnsi 268
OY 301 SVMPAOKITTPAAKYGIPLAYKRYGDKKLHEKKPLOKHQAHOQTPKRVNTEGERRKISE 360
    |||||||
DB 269 svmpaqkltkpaakygiplaykrygdkklhekkplokqhkhqahqtpktrvntgeerrkise 328
OY 361 EAARKRRLEFIEKEKKQ 377
    |||||||
DB 329 eaarkrrlefielckker 345

RESULT 9
AAE06211
ID AAE06211 standard; Protein: 345 AA.
AC AAE06211;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human protein kinase homolog-6 (PKH-6).
XX
KW Human; protein kinase homolog-6; PKH-6; cytosolic; protein therapy;
KW vaccinia; immunosuppressive; antischlerotic; antiabortive; adenocarcinoma;
KW Acquired immune deficiency syndrome; AIDS; melanoma; cancer; bone; liver;
KW breast; autoimmune disorder; multiple sclerosis; drug screening; anaemia;
KW Crohn's disease; ectopic pregnancy; tubal disease; inflammatory disorder;
KW reproductive disorder; polycystic ovary syndrome; asthma.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 4..226
PT /note= "Signature sequence"
XX
PN US6264947-B1.
XX
PD 24-JUL-2001.
XX
PE 20-OCT-1999; 99US-0420915.
XX
PR 15-OCT-1998; 98US-0173581.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Bandman O, Tang YT, Hillman JL, Yue H, Guegler KJ, Corley NC;
PI Gorgone GA, Azimzal Y, Lu DAM;
XX
DR N-PSDB: AAD11848.
XX
DR WPI: 2001-450728/48.
XX
PT Human protein kinase proteins and homologs, useful for preventing,
PT diagnosing and treating cancers, autoimmune/inflammatory disorders and
PT reproductive disorders -

```

```

PS Claim 1: Column 51-54; 38pp; English.
XX
CC The present sequence is human protein kinase homolog-6 (PKH-6).
CC Human protein kinase homologs (PKH) are their cDNA molecules are used in
CC the prevention, diagnosis and treatment of diseases associated with
CC increased or decreased expression of PKH. Examples of such disorders
CC include, cancer (e.g. adenocarcinoma, melanoma and bone, breast and
CC liver cancer), autoimmune/inflammatory disorders (e.g. Acquired immune
CC deficiency syndrome (AIDS), anaemia, asthma, Crohn's disease and
CC multiple sclerosis) and reproductive disorders (e.g. tubal disease,
CC ectopic pregnancy) and polycystic ovary syndrome). PKH, its catalytic or
CC immunogenic fragment are used for screening libraries of compounds in any
CC of the drug screening techniques. PKH nucleic acids are used to generate
CC hybridisation probes useful in mapping the naturally occurring genomic
CC sequences. PKH are also used as antigens in the production of antibodies
CC against protein kinases (PK) and in assays to identify modulators of PK
CC expression and activity. PKH is also used in protein therapy.
XX
SQ Sequence 345 AA;
XX
Query Match 27.8%; Score 1734; DB 22; Length 345;
Best Local Similarity 90.5%; Pred. No. 6.2e-97;
Matches 341; Conservative 3; Mismatches 1; Indels 32; Gaps 1;
OY 1 MEKYVRLQKIGESFGKAILVSTEDGROYVKEINISMSKEREESREAVAVLANMKH 60
    |||||||
DB 1 mekyvrlqkigsgfkgallvkstcdgryvkeinismskereeerravavlanmkh 60
OY 61 PNIVQYRESFENGSLYVMDYCEGDLFKRINAOKGVLFQEDQILDFVQICLAKHVN 120
    |||||||
DB 61 pniyqyresfe-----gllwdfvqiclaikhvn 88
OY 121 DRKILHRDIKSONIFLTRDGTVOLGDFGIARVNSTVELARCTICGTPYLSPEICEKMPY 180
    |||||||
DB 89 drkllhrdiksqnifltrdgtvlgdfgiarvlnstvelarctcigtpylspceicemkpy 148
OY 181 NNSKDIWALGCVLYELCTLKHAFAGSMKNLVKTIISGSPPVSLHYSYDLRSLVSOLF 240
    |||||||
DB 149 nksdiwalgcvlyelctlkhafeagsmknlvklisgsfpvslhysydlrslvsqflk 208
OY 241 RNPDRPSVNSILEKGFIAKRIEFLSPQILAEFCLTKFSKFSQPIPARKRPASGNSI 300
    |||||||
DB 209 rnpdrpsvnsilekgfiakrleflspqilaefcltkfskfsqpiarkrpsagqnsi 268
OY 301 SVMPAOKITTPAAKYGIPLAYKRYGDKKLHEKKPLOKHQAHOQTPKRVNTEGERRKISE 360
    |||||||
DB 269 svmpaqkltkpaakygiplaykrygdkklhekkplokqhkhqahqtpktrvntgeerrkise 328
OY 361 EAARKRRLEFIEKEKKQ 377
    |||||||
DB 329 eaarkrrlefielckker 345

RESULT 10
AAB95221
ID AAB95221 standard; Protein: 320 AA.
AC AAB95221;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:17342.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
DR EP1074617-A2.
XX
DR 07-FEB-2001.
XX
DR 28-JUL-2000; 2000EP-0116126.

```

```
XX 29-JUL-1999: 99JP-0248036.
PR 27-AUG-1999: 99JP-0300253.
PR 11-JAN-2000: 2000JP-0118776.
PR 02-MAY-2000: 2000JP-0183767.
PR 09-JUN-2000: 2000JP-0241899.
XX
XX (HELI-) HELIX RES INSR.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI: 2001-318749/34.
XX
XX primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX
XX Claim 8: SEQ ID 17342: 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and
XX in gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB32446 to
XX AAB95993 represent human amino acid sequences; and AAH13629 to AAH13632
XX represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
XX
XX Sequence 320 AA:
XX
XX Query Match 22.2%; Score 1387; DB 22; Length 320;
XX Best Local Similarity 100.0%; Pred. No. 5.2e-76;
XX Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 208 MKNLVKTIISGSPFVSIHYSDLRSLVQLFRNRPDRPSVSIIEKGFIAKRIEPIIS 267
DB 1 mknlvktisgsfpvslhydsldrslysqldfrnrdpsvsliekfikrlektis 60
OY 268 POLIAEFCILTKESGSPRIAPKRPASGONSISVMPAKITTPAKYGIPLAKYKYGCK 327
DB 61 poliaefcltkesgsopriparkrpasgonsisvmpakittpakyygiplakyygck 120
OY 328 KLIHEKKPIQKHQAHOHPERKRVNTGEERRKRISEBAARKRRRLFEIERKKOKDOITSLMKA 387
DB 121 kihkekpiqkhqahqepkrvntgeerrkriiseaarkrrlfeiekkokqkqdlislmk 180
OY 388 EOMKROEKERLERINRAREOGMRNVNLSAGSGEYKAPFLGSGGTIIPSSFSNGOYEHH 447
DB 181 eqmkroekerlerinrareogmrnvnlsgsgyevkapflgsggtiipssfsngyehh 240
OY 448 AIFDPMQOQRAEDNEAKMKREIYGRGLPER 477
DB 241 aifdpmqqraredneakmkreiygrglper 270
XX
XX RESULT 11
XX " , , , "
```

```
AAU03545
ID AAU03545 standard; Protein; 649 AA..
XX
XX AC AAU03545;
XX
XX DT 12-SEP-2001 (first entry)
XX
XX DE Human protein kinase #45.
XX
XX KW Human; protein kinase; PK; STK; cancer; cardiovascular disease;
XX metabolic disorder; immune related disease; neurological disorder;
XX neurodegenerative disorder; inflammatory disorder; infectious disease;
XX reproductive disorder.
XX
XX OS Homo sapiens.
XX
XX PN WO200138503-A2.
XX
XX PD 31-MAY-2001.
XX
XX PF 22-NOV-2000: 2000WO-US32085.
XX
XX PR 24-NOV-1999: 99US-0167482.
XX
XX PA (SUGEN-) SUGEN INC.
XX
XX PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
XX PI Flanagan P, Clary D;
XX WPI: 2001-343950/36.
XX
XX DR N-PSDB: AAS06745.
XX
XX PT Nucleic acids encoding human kinase polypeptides, useful for preventing
XX diagnosis and/or treating e.g. cancer, immune, cardiovascular and
XX neuronal associated diseases, and microbial infections -
XX
XX XS Claim 7: Figure 2: 433pp: English.
XX
XX AAU03501-AAU03557 represent novel human protein kinases #1-57. The
XX novel protein kinases have been identified as members of the tyrosine
XX or serine/threonine kinase (PK and STK) families. The polynucleotides
XX encoding protein kinases and the polypeptides may be used in the
XX prevention, diagnosis and treatment of diseases associated with
XX inappropriate kinase expression. For example, they may be used to treat
XX cancers (especially cancers of haematopoietic origin), cardiovascular
XX disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
XX immune related diseases (e.g. rheumatoid arthritis), neurological
XX disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
XX Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
XX disease (e.g. HIV) and reproductive disorders (e.g. infertility).
XX Additionally, polynucleotides encoding protein kinases may be
XX used for gene therapy and as DNA probes in diagnostic assays.
XX The protein kinase polypeptides may be used as antigens in the production
XX of antibodies against the protein kinases and in assays to identify
XX modulators of protein kinase expression and activity.
XX
XX Sequence 649 AA:
XX
XX Query Match 16.4%; Score 1025; DB 22; Length 649;
XX Best Local Similarity 32.8%; Pred. No. 9.6e-54;
XX Matches 268; Conservative 125; Mismatches 201; Indels 224; Gaps 25;
OY 1 MEKYVRLOKIGESFKAALIVSTEDGROYVKEINISMSKREKESRREAVLANMKH 60
DB 1 mkyvrvlkaiggaafgkaylakgsdskshvcvikeinfekmpdigeaakkevillemkh 60
OY 61 PNIVQRESEENGSLEYIMDYCEGGDLFKRINACKGVLPQDDLPDFVQICATLKHV 120
DB 61 pnivafnsgengrllfivmcyodgdldmkrlnrgyvlfseddllgfvfqlslhnh 120
OY 121 DRKILHRDIXONIFLTKGTV-QLCDPGIAVLVSLVETLARTCGPYYSPEICMKP 179
DB 121 drkllhrdixonifltkgty-qlcdpgiavlvslvslvartcgpyyspeicmkp 179
```

```

Db 121 drklhhdkaqnlflskngwakiygdfigiarvlnsmelartcigrpylspciqnp 180
QY 180 YNKSQDIWALGCVLYELCTLKAHAFEGSKNKLVLKIIIGSGPPPSLHNSYLSLVSOLF 239
Db 181 ynnktdlwsigcylvelctclhpfegnmqlqvlkicqahhaplspsgrstehslisqll 240
QY 240 KRNRDRPSVNSLLEKGFIAKRIEKLSPQ-----LIAEFLCKTFESK---FGSQPI 288
Db 241 qvsprdrpsinsllkrlpfevfevfwllaeilgcsrlstfgrfrcppl 300
QY 289 -PA-----KRPSGQNSISVMPAKITTKPAARYGIPLAYKKYGDKLHEKPLQKHQA 342
Db 301 hpacimwfkhpiscflqikmlepkaavcghdyayqldmlrrah--kp-----syh 353
QY 343 QTPKRYNTEBERKISEEAAKRRRLFELEKKEKQKDOIISLMAEQMKROEKERLERIN 402
Db 354 pldge--ntgve-----dygqetrhgsp--sqwpaeylqrfesqyqlk 395
QY 403 RARQGMNRYLSAGSGEVKAPPLGSGGTIAPSSFSRCQYBHYHAIFDQMOQRAEDNE 462
Db 396 vekqlglr-----psaeapn-----ynqrqelrnsgee 423
QY 463 AKMKREITYRGRLPEPROKGLAVERAKOVEEFLQRRKEAMQKARAKGHMYLARLRQRL 522
Db 424 prlge-----lprfkemkeqeywqlge-----llrq 450
QY 523 QNFENRQIRAKKLGEKKEANHSEGGSGSEADMRRKKTIESLKAANRAAVKEQLERK 582
Db 451 qyhdmkeikrlkmgrpedl-----eklikmrirgntk-----esk 486
QY 583 RKEAVEREKKWEHEHLVAKGVKSSDVSPPLOGHETGSGSPSKOQMSVLSVTSALKEVGD 642
Db 487 npe-----qkykakkgvkt-----ealnd 505
QY 643 SLIDTRETSEMQTNNAISKRREILRLNENLKAQDEKGMQNLSTFE---INVHE 698
Db 506 kciisd-----enllqeeamdiptellfedgmkfeye 539
QY 699 DAKEH-----EK-----EKSYS-----DRKKWEAGG-QLYVILDELTLDTSTFT 738
Db 540 cvkehgdyltkafekllhpcpegfstqlvaavgnrrgwdgagpqlqlgmaavadlscpt 599
QY 739 ERHTGEVYIKLGPNGSPRAWAGKSPDTSVKLIIGEAEL 776
Db 600 gpdn-gqylvleglpgnrtkqwrheapptlmsvlaaahl 636

RESULT 12
AAM78344
ID AAM78344 standard; Protein; 506 AA.
XX
AC AAM78344;
XX
XX 06-NOV-2001 (first entry)
DE Human protein SEQ ID NO 1006.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
OS Homo sapiens.
XX
XX WO200157190-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US04098.
XX
XX 03-FEB-2000; 2000US-0496914.
XX
XX 27-APR-2000; 2000US-0560875.
XX
XX 20-JUN-2000; 2000US-0598075.

```

```

PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
XX (HSE-) HYSQ INC.
XX
XX Tang YF, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
XX Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
XX Xue AY, Yang Y, Wejhrman T, Goodrich R;
XX WPI: 2001-476283/51.
XX N-PSDB: AAK51477.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
XX useful in diagnosis and gene therapy -
XX
XX Claim 20; Page 3232-3233; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
XX (AAM80020) are omitted as the relevant pages from the sequence listing
XX were missing at the time of publication.
XX
XX Sequence 506 AA:
XX
XX Query Match 14.18; Score 882; DB 22; Length 506;
XX Best Local Similarity 42.58; Pred. No. 3e-45;
XX Matches 175; Conservative 70; Mismatches 103; Indels 64; Gaps 6;
XX
XX 1 MEKYYRLQKIGSGFGKNAIIVKSTEDGRQYVIREINISMSSKERESREYAVIANKKH 60
XX 1 mddymvlmliqegsfgrallvqhesnqmfamkelrlpk-sfsntqnsrkeavvllakmh 59
XX
XX 61 PNIVQRESFEENGSLYVMDYCEGDFRINAQKGVLFQEDQIIDWFQICLAKHVN 120
XX 60 pvlvafkesfeeghlylmeycdggdlmqkikqkgkllfpedmlnwftqmclygmnhh 119
XX
XX 121 DRKILHROIKSONIFLTKDGVQGLDGFARVNSVTELAARCICIGPYLSPSEICENRY 180
XX 120 ktrvlhrdksknflitqngkvklygdfgsarllsnpmafacylvgyppvypveiwenlpy 179
XX
XX 181 NKRSDIWMALGCVLYELCTLKAHAFEGSKNKLVLKIIIGSGPPPSLHNSYLSLVSOLF 240
XX 180 mksdtdlwsigcylvelctclhpfegnmqlqvlkicqahhaplspsgrstehslisqll 239
XX
XX 241 RNRDRPSVNSLLEKGFIAKRIEKLSPOLIAEFLCKTFESKSGSPRIAKRPSAGQNSI 300
XX 240 rnpshpsactcllstrgylavtqclpdelme----- 272
XX
XX 301 SYMPAKITTKPAARYGIPLAYKKYGDKLHEKPLQKHQAQOTPEKRYNTEBERKISE 360
XX 273 -----ygeevleelk-----nshncprrkknpcstfitalign 304
XX
XX 361 EAARKRRLFELEKKEKQKDOIISLMAEQMKROEKERLERINARQGMNRY 412
XX 305 eastvg-----eeegdrkgshtdlesine--nlvesalltrvr-eeekgnksv 348

RESULT 13
AAM79328

```

ID AAM79328 standard; Protein: 527 AA.
 AC AAM79328;
 DT 06-NOV-2001 (first entry)
 DE Human protein SEQ ID NO 2974.
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX Homo sapiens.
 OS Mo200157190-A2.
 PN 09-AUG-2001.
 PD 05-FEB-2001; 2001MO-US04098.
 PE 03-FEB-2000; 2000US-046914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0653561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y,
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wehrman T, Goodrich R;
 DR WPI: 2001-476283/51.
 DR N-PSDB: AAK52461.
 XX Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 PS Claim 20: Page 218; 6221pp; English.
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX Sequence 527 AA:

Query Match 14.0%; Score 875; DB 22; Length 527;
 Best Local Similarity 42.2%; Pred. No. 8.4e-45;
 Matches 174; Conservative 70; Mismatches 104; Indels 64; Gaps 6;

QY 1 MEKYLKIGESFGKALIVKSTEDGROVIVKEINSRMSKRESESRVAVLANMKH 60
 DB 22 mddyvmvltmigsfgralllvghessngmfamkelrlpk-sfentqnskeavllakmh 80
 QY 61 ENIVYRSEFENGSLTYVMDYCGSDLPKRNAGGVLPQEQIILDMVQICLAKNHH 120
 DB 81 pniyafkesfaeghlylvmeycdgdlmkikqgkylfpedmlinwftqmclyvnhn 140

QY 121 DRKILHRDIKSONIFLTQDGTVOJLDFGIARVLNSTVELARFCIGTPYVLSPEICENKPY 180
 DB 141 kkrvlhrdtksknifltqngkglqdfgsarlspmafactyvgfpyvpelwani 200
 QY 181 NNRSDIWAIGCVLYELCTKHAFAEAGSMKNLVKLIISGFPPVSLHYSYDLRSVLQLF 240
 DB 201 nksdswslgcllyelctkhpfganswnkllkvcgcslpsbhsyqlfvlxqmfk 260
 QY 241 RNRDRPPSVSILKEKFIKRIEFLSPOLIAEEFCCLKTFKFCGSPDIPAKRPASGONSI 300
 DB 261 rnpshrsatlslrgslvarlvqclpelpme----- 293
 QY 301 SVMPAQKITRPAKVCIGPLAYKKYGDKKLHEKKPLQKKHQAHPKRVNTGEERRKISE 360
 DB 294 -----ygeevleek-----nshhnprrkktprtrialgn 325
 QY 361 EAARKRLEFIEKKOKDQIISLMAKAGMKRQEKERLERINRAREGGRNV 412
 DB 326 eastvg-----eeqdrkshdlesine--nlvesalrrvr-eeqynksv 369
 RESULT 14
 ID AAY68778 standard; Protein: 510 AA.
 AC AAY68778;
 DT 16-MAY-2000 (first entry)
 DE Amino acid sequence of a human phosphorylation effector PHSP-10.
 XX Human; phosphorylation effector; PHSP; proliferative disorder;
 KW immune disorder; neuronal disorder.
 XX Homo sapiens.
 OS
 PA Key Location/Qualifiers
 PI Modified-site 24 /note= "potential phosphorylation site"
 PI Modified-site 51 /note= "potential phosphorylation site"
 PI Region 52..261 /note= "protein kinase family signature sequence"
 FT Modified-site 152
 FT Modified-site 185 /note= "potential phosphorylation site"
 FT Modified-site 201 /note= "potential glycosylation site"
 FT Modified-site 210 /note= "potential phosphorylation site"
 FT Modified-site 247 /note= "potential phosphorylation site"
 FT Modified-site 292 /note= "potential phosphorylation site"
 FT Modified-site 297 /note= "potential phosphorylation site"
 FT Modified-site 312 /note= "potential phosphorylation site"
 FT Modified-site 323 /note= "potential phosphorylation site"
 FT Modified-site 325 /note= "potential phosphorylation site"
 FT Modified-site 329 /note= "potential phosphorylation site"
 FT Modified-site 349 /note= "potential phosphorylation site"
 FT Modified-site 358 /note= "potential glycosylation site"
 FT Modified-site 377 /note= "potential phosphorylation site"
 FT Modified-site 381 /note= "potential phosphorylation site"

Wed May 15 14:13:54 2002

us-09-783-320-4.rag

Page 14

Sequence 698 AA:

Search completed: May 15, 2002, 07:52:53
Job time: 4931 sec

Query Match 10.1% Score 628; DB 22; Length 698;
Best Local Similarity 23.8% Pred No. 1e-29;
Matches 227; Conservative 116; Mismatches 271; Indels 340; Gaps 29;

QY 1 MEKYVRLQKIGESFGKALIVKSTEDGROYIKEINISMSKREBSREAVLANMKH 60
DB 1 mekyvrlvgrgavghicrlkadqvlvklqivgcmkceerqaqecqvlklhnh 60
QY 61 PNIVQRESFEENGSLYVMDYCEGDLFKRINAKGVLFQEDQILDMVQICLAKHVN 120
DB 61 pnivleyenflfedkalmameyapqglaeifqrcnslleetiilhfvtqllalhvh 120
QY 121 DRKILARDIKSONIFLTKDG-TVQLGDGIARVLNSTVELARTCIGTPYVLSPEICENP 179
DB 121 thlllhrdlkqnllldkhrmvkigdfgisklisk-----stpcyispelceqxp 172
QY 180 YNNKSDIWLGCVLEYELCTLKAFAEGSMKNLYLKISGFPVSLHYSDRLSVQLF 239
DB 173 yngkediwalgvlyelaslkrateaanlpalvikimgtlfaplsdryspeqlvlsll 232
QY 240 KNPDRPVSNTLEKGFIAKRIEFLSPOLIAEFCLTKTFKSGSPPIPAKRPAAGONS 299
DB 233 sljepagrpplshlm-----agplclral----- 255
QY 300 ISVMAQKITKPAKYGIPLAKYKGDKLIHE-----KKPLQKHQAHOPEKRVNTGEE 354
DB 256 -----lnltgdrevrpgqnrqdnqcpdq----- 281
QY 355 RKKISEBARRRLEFIEKEKKOKDQIISLMAKADQMRKOEKERLERINARQGMNVLS 414
DB 282 -----rgllm 286
QY 415 AGGSEVKAPFLGSGGTIAPSSFGQYEHYNAIFDOMOQORAD----- 460
DB 287 lfgs-----gsngcighslldisqptlvealigyemvqveealstllgsapldq 338
QY 461 -----NEAKMKREIYGRG-----LPERQKGL--AVERAKOVEEF--LQRRK 498
DB 339 epllsldigtahsaavtgeedlgsqdvnrllpswerghllagvaastdvstfsegdckepd 398
QY 499 EAMQKKARAECHMY-----LARLROIRLQNFNEROQIKAKLGEKKEANHSEG-- 547
DB 399 kecwrhkgctghllypfasdcvrhslhshvnhcnstr-----lkdsedssssrgag 452
QY 548 -----OEGSEEDMRKKITSLKANANARAVALKEOL-----ERKRKE 585
DB 453 ptcshvlespcfeltpreeehverfrlygwksyrvsvavlnhpllyhecgaddlmkkkkr 512
QY 586 AVEREKKWEHLVAKYKSDVDPPLGOHETGSGPSKQOMRSYISVTSALKEVGDSSI 645
DB 513 rrrkxkpp-----lptqvgpataspdlgtmatgtp-----dsta 547
QY 646 TPTRETSEMOXTNNAISSKREILRLNENLKAQEDKGMQNLSDTFEINVHEDAKEHEK 705
DB 548 pltiwrse--sptkqgqsk--vikkykkkkekckdeemd-----ekaklkxk 592
QY 706 EKSVSSDRKKWEGCOLVPLDELTLDTSFSTRHTVGEVIRKIPNGSPRRAMGKSPD 765
DB 593 akkgqltkkk-----spvklep-----spdp 613
QY 766 SVLKILGEAELOLOTELENTTIRSEISPEGEKYKPLITGEKVVOCISHEINPSALVDS 825
DB 614 -vsrlsarqlarmse--sspesreleese-day-----ngrggelsedl----- 656
QY 826 VETKSPERSE-----ASPQMSIKESGLEPPDLETEIIOEPGCTNKDES 870
DB 657 veassprkrentvgakktgaktgskqarkvnr-----kspgsgnplns 698

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (COPY)